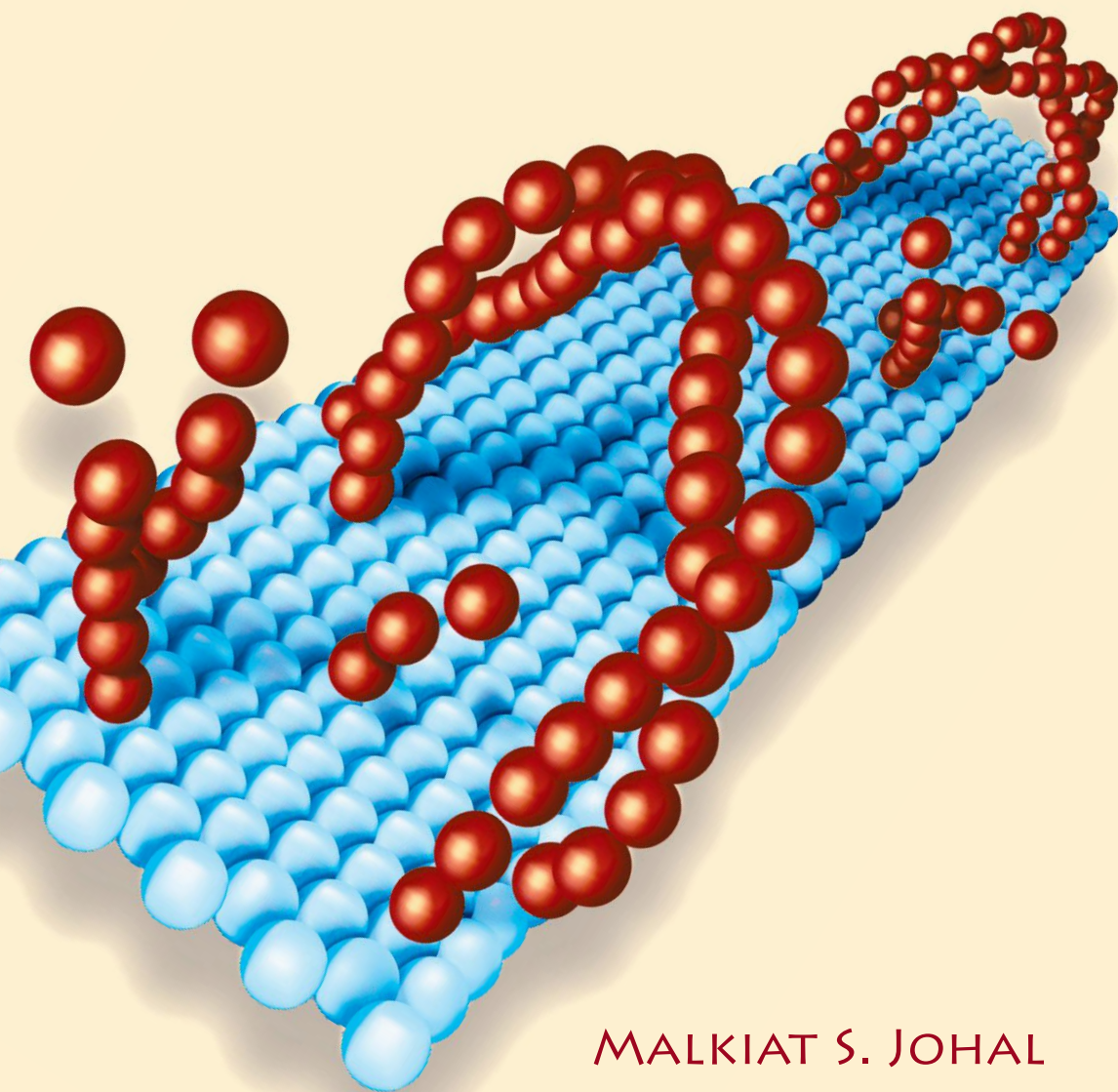


# UNDERSTANDING NANOMATERIALS



MALKIAT S. JOHAL



CRC Press  
Taylor & Francis Group

# UNDERSTANDING NANOMATERIALS



# UNDERSTANDING NANOMATERIALS

MALKIAT S. JOHAL



CRC Press

Taylor & Francis Group

Boca Raton London New York

---

CRC Press is an imprint of the  
Taylor & Francis Group, an **informa** business

Front cover image was provided by Dr. Gerry Ronan, Farfield Group, Manchester, United Kingdom.

CRC Press  
Taylor & Francis Group  
6000 Broken Sound Parkway NW, Suite 300  
Boca Raton, FL 33487-2742

© 2011 by Taylor & Francis Group, LLC  
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works  
Version Date: 20110720

International Standard Book Number-13: 978-1-4398-9112-4 (eBook - PDF)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access [www.copyright.com](http://www.copyright.com) (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

**Trademark Notice:** Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

**Visit the Taylor & Francis Web site at**  
**<http://www.taylorandfrancis.com>**

**and the CRC Press Web site at**  
**<http://www.crcpress.com>**

*This book is dedicated to my  
beautiful daughter Simran*



# Contents

---

Preface . . . . .	xv
Acknowledgments . . . . .	xvii
About the Author . . . . .	xix

<b>Chapter 1</b>	A Brief Introduction to Nanoscience . . . . .	1
1.1	The Need for Nanoscience Education. . . . .	1
1.2	The Nanoscale Dimension and the Scope of Nanoscience . . . . .	2
1.3	Self-Assembly. . . . .	5
1.4	Supramolecular Science . . . . .	6
1.5	Sources of Information on Nanoscience . . . . .	7

<b>Chapter 2</b>	Intermolecular Interactions and Self-Assembly. . . . .	9
	Chapter Overview . . . . .	9
2.1	Intermolecular Forces and Self-Assembly. . . . .	9
2.1.1	Ion-Ion Interactions . . . . .	11
2.1.2	Ion-Dipole Interactions . . . . .	13
2.1.3	Dipole-Dipole Interactions . . . . .	16
2.1.4	Interactions Involving Induced Dipoles . . . . .	17
2.1.5	Dispersion Forces. . . . .	20
2.1.6	Overlap Repulsion . . . . .	21
2.1.7	Total Intermolecular Potentials . . . . .	24
2.1.8	Hydrogen Bonds. . . . .	25
2.1.9	The Hydrophobic Effect. . . . .	27
2.2	Electrostatic Forces Between Surfaces: The Electrical Double Layer . . . . .	28
2.2.1	The Electrical Double Layer . . . . .	28
2.2.2	The Debye Length . . . . .	31
2.2.3	Interactions Between Charged Surfaces in a Liquid. . . . .	32
2.3	Intermolecular Forces and Aggregation. . . . .	34
2.4	Simple Models Describing Electronic Structure . . . . .	36



2.4.1	The Particle in a Box Model . . . . .	37
2.4.2	Conjugation in Organic Molecules . . . . .	43
2.4.3	Aggregation and Electronic Structure . . . . .	47
2.4.4	$\pi$ - $\pi$ Stacking Interactions. . . . .	49
	References and Recommended Reading . . . . .	49
	End of Chapter Questions . . . . .	51
<b>Chapter 3</b>	<b>Rudiments of Surface Nanoscience . . . . .</b>	<b>55</b>
	Chapter Overview . . . . .	55
3.1	Fundamentals of Surface Science . . . . .	55
3.1.1	The Surface Energy of Solids and Liquids . . . . .	55
3.1.2	Surface Free Energy of Adsorbed Monolayers . . . . .	56
3.1.3	Contact Angles and Wetting Phenomena. . . . .	61
3.1.4	Nanomaterials and Superhydrophobic Surfaces . . . . .	62
3.2	Adsorption Phenomena: Self- Assembled Monolayers . . . . .	68
3.2.1	Simple Adsorption Isotherms. . . . .	73
3.2.2	Other Useful Adsorption Isotherms . . . . .	78
3.3	Surfactant Chemistry . . . . .	79
3.3.1	Micelle and Microemulsion Formation. . . . .	82
3.3.2	The Determination of Surface Excess: The CMC and the Cross- Sectional Area per Molecule. . . . .	84
	References and Recommended Reading . . . . .	88
	End of Chapter Questions . . . . .	89
<b>Chapter 4</b>	<b>Characterization at the Nanoscale . . . . .</b>	<b>97</b>
	Chapter Overview . . . . .	97
4.1	Surface Tensiometry: The Surface Tensiometer . . . . .	97
4.2	Quartz Crystal Microbalance . . . . .	101
4.2.1	The Piezoelectric Effect. . . . .	102
4.2.2	QCM Principles . . . . .	103
4.2.3	QCM and Dissipation (D). . . . .	105
4.2.4	Modern QCM-D Setup . . . . .	107

4.3	Ellipsometry. . . . .	108
4.3.1	Basic Principles of Electromagnetic Theory and Polarized Light. . . . .	109
4.3.2	Basic Principles of Ellipsometry. . . . .	113
4.3.3	Obtaining the Thickness of Films: Optical Parameters $\Delta$ and $\Psi$ ( $\psi$ ) . . . . .	115
4.3.4	The Ellipsometer . . . . .	117
4.4	Surface Plasmon Resonance . . . . .	119
4.4.1	Principles of SPR . . . . .	120
4.4.2	SPR Instrument Setup . . . . .	122
4.5	Dual Polarization Interferometry . . . . .	124
4.5.1	Waveguide Basics. . . . .	124
4.5.2	Waveguide Interferometry and the Effective Refractive Index. . . . .	125
4.5.3	Principles of Dual Polarization Interferometry . . . . .	128
4.5.4	Parameters of a DPI Instrument and Common Applications . . . . .	130
4.6	Spectroscopic Methods . . . . .	130
4.6.1	Interactions Between Light and Matter . . . . .	132
4.6.2	UV-Visible Spectroscopy. . . . .	135
4.6.2.1	Principles of UV-Visible Spectroscopy . . . . .	135
4.6.2.2	Setup of a UV-Visible Spectrophotometer. . . . .	136
4.6.3	The Absorption of Visible Light by a Nanofilm. . . . .	138
4.6.4	Molecular Fluorescence Spectroscopy . . . . .	141
4.6.4.1	Principles of Fluorescence and Fluorescence Quantum Yield . . . . .	141
4.6.4.2	Setup of a Fluorometer for Bulk Phase and Thin Film Fluorescence Measurements . . . . .	143
4.6.5	Vibrational Spectroscopy Methods . . . . .	144
4.6.5.1	Introduction to Vibrational Modes . . . . .	144

	4.6.5.2	Attenuated Total Reflection IR Spectroscopy . . . . .	148
	4.6.5.3	Reflection Absorption IR Spectroscopy . . . . .	150
	4.6.6	Raman Spectroscopy . . . . .	150
	4.6.6.1	Rayleigh and Raman Light Scattering . . . . .	150
	4.6.6.2	Surface Enhanced Raman Spectroscopy . . . . .	154
4.7		Nonlinear Spectroscopic Methods . . . . .	155
	4.7.1	An Introduction to Nonlinear Optics . . . . .	155
	4.7.2	Second-Harmonic Generation . . . . .	159
	4.7.3	Sum-Frequency Generation Spectroscopy . . . . .	162
4.8		X-Ray Spectroscopy . . . . .	167
	4.8.1	Absorption . . . . .	167
	4.8.2	Fluorescence . . . . .	167
	4.8.3	Diffraction . . . . .	168
4.9		Imaging Nanostructures . . . . .	169
	4.9.1	Imaging Ellipsometry . . . . .	169
	4.9.1.1	Imaging Using Conventional Ellipsometry . . . . .	169
	4.9.1.2	Principles of Modern Imaging Ellipsometry . . . . .	169
	4.9.1.3	Methods for Extracting Ellipsometric Data in Imaging Ellipsometry . . . . .	170
	4.9.1.4	Image Focusing . . . . .	171
	4.9.1.5	Resolution of an Imaging Ellipsometer . . . . .	171
	4.9.2	Scanning Probe Methods . . . . .	172
	4.9.2.1	Scanning Tunneling Microscopy . . . . .	172
	4.9.2.2	Atomic Force Microscopy . . . . .	176
	4.9.3	Transmission Electron Microscopy . . . . .	179
	4.9.3.1	Principles of TEM . . . . .	179
	4.9.3.2	TEM Instrumentation . . . . .	183
	4.9.4	Near-Field Scanning Optical Microscopy . . . . .	185

4.9.4.1	History and Principles of NSOM . . . . .	186
4.9.4.2	Modern NSOM Instrumentation and Different NSOM Operating Modes . . . . .	187
4.10	Light Scattering Methods . . . . .	191
4.10.1	The Measurement of Scattered Light: Determining the Aggregation Number of Micelles . . . . .	192
4.10.2	Dynamic Light Scattering . . . . .	195
	References and Recommended Reading . . . . .	200
	End of Chapter Questions . . . . .	200
<b>Chapter 5</b>	<b>Types and Uses of Some Nanomaterials . . . . .</b>	<b>207</b>
	Chapter Overview . . . . .	207
5.1	Supramolecular Machines . . . . .	207
5.1.1	Model Dye System . . . . .	208
5.1.2	Photorelaxation . . . . .	211
5.1.3	Formation and Properties of the Exciton . . . . .	212
5.2	Nanowires . . . . .	215
5.2.1	Basic Quantum Mechanics of Nanowires . . . . .	216
5.2.2	Conductivity . . . . .	217
5.2.3	Nanowire Synthesis . . . . .	218
5.2.4	Summary . . . . .	219
5.3	Carbon Nanotubes . . . . .	220
5.3.1	Carbon Nanotube Structure . . . . .	220
5.3.2	Some Properties of Nanotubes . . . . .	222
5.3.3	Methods for Growing Nanotubes . . . . .	222
5.3.3.1	Arc Discharge . . . . .	223
5.3.3.2	Laser Ablation . . . . .	223
5.3.3.3	Chemical Vapor Deposition . . . . .	224
5.3.4	Catalyst-Induced Growth Mechanism . . . . .	224
5.4	Quantum Dots . . . . .	225
5.4.1	Optical Properties . . . . .	226
5.4.2	Synthesis of Quantum Dots . . . . .	227
5.4.2.1	Precipitative Methods . . . . .	227

	5.4.2.2	Reactive Methods in High-Boiling-Point Solvents . . . . .	228
	5.4.2.3	Gas-Phase Synthesis of Semiconductor Nanoparticles . . . . .	229
	5.4.2.4	Synthesis in a Structured Medium . . . . .	229
	5.4.3	In Vivo Molecular and Cell Imaging . . . . .	229
5.5		Langmuir-Blodgett Films . . . . .	230
	5.5.1	Langmuir Films . . . . .	232
	5.5.2	Langmuir-Blodgett Films . . . . .	234
5.6		Polyelectrolytes . . . . .	237
	5.6.1	Electrostatic Self-Assembly . . . . .	238
	5.6.2	Charge Reversal and Interpenetration . . . . .	241
	5.6.3	Multilayer Formation . . . . .	243
5.7		Model Phospholipid Bilayer Formation and Characterization . . . . .	244
	5.7.1	Black Lipid Membranes . . . . .	245
	5.7.2	Solid Supported Lipid Bilayers . . . . .	246
	5.7.3	Polymer Cushioned Phospholipid Bilayers . . . . .	250
	5.7.4	Fluorescence Recovery after Photobleaching . . . . .	250
	5.7.5	Fluorescence Resonant Energy Transfer . . . . .	252
	5.7.6	Fluorescence Interference Contrast Microscopy . . . . .	253
5.8		Self-Assembled Monolayers . . . . .	255
	5.8.1	Thiols on Gold . . . . .	256
	5.8.2	Silanes on Glass . . . . .	257
5.9		Patterning . . . . .	258
	5.9.1	Optical Lithography . . . . .	259
	5.9.2	Soft Lithography . . . . .	262
	5.9.3	Nanosphere Lithography . . . . .	262
	5.9.4	Patterning Using AFM . . . . .	263
	5.9.5	Summary . . . . .	266
5.10		DNA and Lipid Microarrays . . . . .	266
	5.10.1	Using a DNA Microarray . . . . .	266
	5.10.2	Array Fabrication . . . . .	267
	5.10.3	Optimization . . . . .	270
	5.10.4	Applications . . . . .	271

5.10.5 Arrays of Supported Bilayers and Microfluidic Platforms . . . . .	271
5.10.6 Summary . . . . .	274
Cited References . . . . .	274
References and Recommended Reading . . . . .	276
End of Chapter Questions . . . . .	278
Appendix . . . . .	281
Glossary . . . . .	283



# Preface

---

## TO THE STUDENT

Nanoscience is a rapidly changing field where new innovations and discoveries are being made every day. To write a book that captures even a fraction of what the scientific literature has produced over the last few years would be a monumental task. With this in mind, the topics presented in this book are carefully selected to provide a basic understanding of the field. Many important topics, such as computational chemistry and solid state physics, have been given limited coverage, largely because I want this book to be accessible to any student who has taken introductory college-level science courses.

This book is written for a full- or half-semester course in nanoscience with an emphasis on understanding nanomaterials. The stress on “understanding” is the key behind the objective of the text: to provide fundamental insight into the molecular driving force underlying self-assembly processes, as well as to explain how to characterize the resulting nanomaterials. Knowledge of self-assembly and characterization is essential for an understanding of these interesting systems.

It should be noted that this book does not draw heavily from scientific papers; rather, it should be used in conjunction with the primary literature.

## TO THE INSTRUCTOR

I have drawn relevant material from many scientific disciplines, assuming only a basic level of competency in physics, chemistry, and biology. Mathematical rigor has been limited to presenting key results and simple proofs. Instructors should use their discretion in placing emphasis on or “filling holes” in areas that may seem somewhat inadequate or limited in scope. The half-course model is suggested for teaching material directly from the book and solving the end-of-chapter problems. For a full-semester course, the book should be used in a course that requires students to refer to the primary literature. The latter may be more suitable for intermediate to



advanced-level classes, although I strongly suggest training students to read papers early in their careers.

The approach taken in this book is to focus on preparing the student to read papers in this field, and so I have limited specific examples to landmark papers. This book should provide the necessary background to enable the student to comprehend articles from scientific journals. I teach the material from this book in conjunction with a student seminar series in which the students select interesting papers for presentation and class discussion. This approach combines discussion-based and problem-solving skills, and provides exposure to a highly interdisciplinary field of study.

# Acknowledgments

---

Writing this book would not have been possible without the support of friends and family. My research students and members of my Soft Nanomaterials class (2008 and 2009) have been crucial in helping me to develop a textbook that meets the needs of students interested in this field. I would like to express sincere gratitude to Robert Rawle, Theodore Zwang, Michael Haber, Michael Gormally, and Thomas Lane for making valuable contributions, including editing and extensive proofreading of the manuscript. In addition, for their detailed review of the manuscript and constructive comments, I'm very grateful to Professors Lisa Klein at Rutgers University, Joseph Tracy at North Carolina State University, and Marcus Lay at University of Georgia. Last but not least I would like to thank my editor, Luna Han, for her invaluable suggestions and her unwavering commitment. This book would not have been written without her dedication.

**Malkiat S. Johal**

*Claremont, California*

*malkiat.johal@pomona.edu*



# About the Author

---



**Malkiat S. Johal** is Associate Professor of Physical Chemistry at Pomona College, Claremont, California. He obtained a first-class honors degree in chemistry from the University of Warwick, UK. After earning his Ph.D. in physical chemistry from the University of Cambridge, UK, Dr. Johal joined Los Alamos National Laboratory, New Mexico, as a post-doctoral research associate, where he worked on the nonlinear optical properties of nanoassemblies. His research laboratory at Pomona College focuses on using self-assembly and ionic adsorption processes

to fabricate nanomaterials for optical and biochemical applications. Professor Johal's laboratory also explores fundamental phenomena such as ion-pair complexation, adsorption, surface wettability, and intermolecular non-covalent interactions that lead to highly ordered nanostructures. He teaches courses in physical chemistry and soft nanomaterials. *Understanding Nanomaterials* is his first book.



## A Brief Introduction to Nanoscience

---

### 1.1 THE NEED FOR NANOSCIENCE EDUCATION

Significant growth in areas of scientific research inevitably impacts education, first at the graduate level, and then at the undergraduate level. In the last decade a healthy flow of publications has addressed the needs of graduate students and trained professionals in the field of nanoscience. Rapid research advances in areas such as soft matter, supramolecular science, and biophysical chemistry have fueled the recent surge in the number of professional journals in nanotechnology. Over 50 journals are publishing research in nanotechnology, and over 30 are devoted solely to nanomaterials.

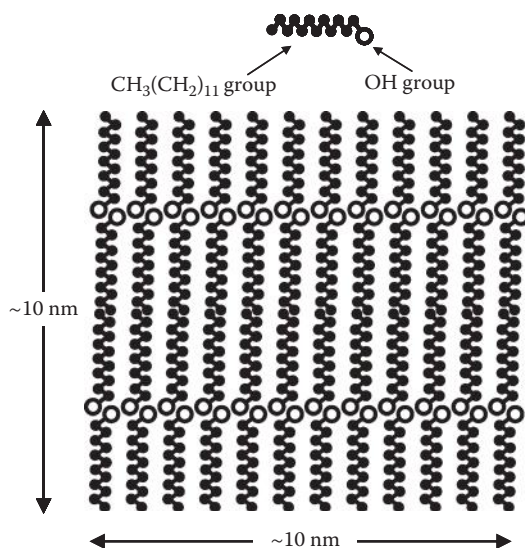
Examples of new, devoted publications include the Royal Society of Chemistry's *Soft Matter* and *Nanoscale*, which cover the interdisciplinary science underpinning the properties and applications of soft matter at the nanoscale, as well as the American Chemical Society's *ACS Nano* and *Nano Letters*. In addition, there has been steady growth in funding for nanoscience from both private sources and government agencies such as the National Science Foundation, Department of Energy, and National Institute of Standards and Technology. Together with the growth in research activity and sharp increase in professional publications, this rise in funding commitment provides a compelling reason to begin serious training of our future workforce in this area. Thus, the present textbook has been developed with the needs of the undergraduate curriculum in mind. This introductory textbook encompasses the fundamental principles governing the fabrication of nanomaterials, processes such as self-assembly, patterning, and nanolithography. Significant coverage is also given to the characterization and applications of such materials. The book captures the interdisciplinary nature of this field, and attempts to provide a well-balanced approach to teaching nanoscience (although

much of the material emphasizes liquids and surface science). The book also showcases nanomaterials research originating from the chemistry, biology, physics, medicine, and engineering communities. In particular the book emphasizes, wherever possible, topics of current global interest (energy, the environment, and medicine). The target audience is undergraduates with at least one year of introductory chemistry, physics, and (preferably) biology.

## 1.2 THE NANOSCALE DIMENSION AND THE SCOPE OF NANOSCIENCE

The word “nano” is derived from the Latin word *nanus*, meaning “dwarf,” and is often used in the context of miniaturization. It is given the abbreviation n. In the international systems of units, nano is the prefix used when multiplying a unit, such as a given length, by  $10^{-9}$ . Thus, one can speak of a nanometer ( $1 \text{ nm} = 1 \times 10^{-9} \text{ m}$ ), a nanosecond ( $1 \text{ ns} = 1 \times 10^{-9} \text{ s}$ ), and even a nanogram ( $1 \text{ ng} = 1 \times 10^{-9} \text{ g}$ ). The nano term is typically used to refer to objects with length scale approaching the order of  $10^{-9} \text{ m}$ . Thus, one can speak of nanotubes, nanofossils, nanowires, and nanofilms as materials in which at least one dimension is on the order of  $10^{-9} \text{ m}$ . The term nanofabrication refers to the procedure used to construct materials with nanoscale dimension.

To put the nanoscale in perspective, consider the size of a hydrogen atom. You may have learned from introductory physics or chemistry that the Bohr radius (the distance from the 1s electron to the central proton in hydrogen) is about 52.0 pm, or roughly 0.05 nm. This distance arguably represents the lowest limit with respect to atomic distances. In fact, atoms and their ions vary in size between this number and about 0.3 nm. This range represents the atomic scale. The hydrogen molecule,  $\text{H}_2$ , has a proton-proton distance (bond length) of about 0.07 nm. The much larger  $\text{I}_2$  molecule has a bond length of about 0.3 nm, and the diameter of the benzene ring is about 0.5 nm. The size of molecules increases rapidly with structural complexity. The polyatomic molecule dodecanol,  $\text{CH}_3(\text{CH}_2)_{10}\text{CH}_2\text{OH}$ , has a length approaching 2 nm. If 48 such molecules were stacked together as shown in Figure 1.1, then it is conceivable that an aggregate of length  $\sim 10 \text{ nm}$  and height  $\sim 10 \text{ nm}$  could be formed. Nanostructures may be comprised of thousands of such molecules resulting in aggregates on the scale of hundreds of nanometers. Furthermore, macromolecular systems



**Figure 1.1** An assembly of 48 dodecanol molecules forming a hypothetical aggregate of length and width 10 nm. The filled circles represent carbon atoms, and the open circle represents the oxygen of the terminal hydroxyl group of the molecule. Hydrogen atoms are not shown for convenience. This two-dimensional assembly is held together by a combination of hydrogen bonds and hydrophobic interactions (discussed further in Chapters 2 and 3).

such as polymers and proteins have sizes approaching 10 nm; the aggregation of such molecules may result in structures on the scale of micrometers.

There is no clear boundary between what one considers molecular scale and the nanoscale that characterizes aggregated systems. It is simply a matter of the relative size of the molecule making up the aggregate and the size of the aggregate itself. It is sufficient to say that the field of nanotechnology or nanoscience deals with the manipulation and control of structures of a length scale 1000 nm or smaller. The science is fascinating because physical and chemical phenomena at these scales are markedly different from those observed in bulk (macroscopic) matter. Sometimes the difference is just a result of the much larger surface-area-to-volume ratio as particles shrink in size. Therefore, surface science plays a central role in understanding nanomaterials.

Nanoscience is full of possibilities and presents us with the potential for significant technological breakthroughs in the near future. Nobel laureate Richard P. Feynman realized the importance of this field almost six decades ago. In



his legendary speech, “There’s Plenty of Room at the Bottom,” he stated: “This field is not quite the same as others in that it will not tell us much of fundamental physics in the sense of, ‘What are the strange particles?’ But it is more like solid state physics in the sense that it might tell us much of great interest about strange phenomena that occur in complex situations. Furthermore, a point that is most important is that it would have an enormous number of technical applications” (see <http://www.zyvex.com/nanotech/feynman.html>).

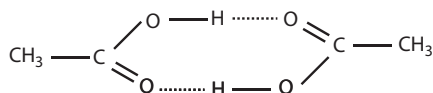
Nanotechnology is defined as the engineering or manipulation of functional systems at the molecular scale. A functional system is used to describe a material that has a well-defined responsibility and performs that responsibility with little or no “side-effects.” Although the term nanotechnology was popularized in the 1980s, scientists have been studying nanostructures for well over a century. As early as the mid-1800s, Michael Faraday investigated the properties of colloidal gold and attempted to explain the ruby color of such solutions. These solutions were made by treating an aqueous solution of  $\text{NaAuCl}_4$  with a reducing agent. Faraday concluded that the resulting ruby-colored fluid was very finely divided gold dispersed in the aqueous solution. A century later, electron microscopy showed that these particles were indeed gold particles with average diameters of around 6 nm.

Over the last few decades, nanotechnology has focused largely on the use of colloidal systems and simple nanoparticles in coatings and materials. For example, silver nanoparticles have found use in hundreds of products because of their antimicrobial properties. More recently, nanotechnology has been used to explore biologically active materials as novel biosensors and targeted drug delivery vehicles for the treatment of diseases. The field is also impacting microelectronics, or more appropriately nanoelectronics, with the development of new transistors, amplifiers, and adaptive structures. The next few decades will inevitably move nanotechnology to the point where we will be able to create complex nanosystems and molecular devices by design. Richard Feynman envisioned the capability of nanotechnology over 50 years ago stating, “I want to build a billion tiny factories, models of each other, which are manufacturing simultaneously. ... The principles of physics, as far as I can see, do not speak against the possibility of maneuvering things atom by atom. It is not an attempt to violate any laws; it is something, in principle, that can be done; but in practice, it has not been done because we are too big.”

### 1.3 SELF-ASSEMBLY

In general there are two strategies for designing molecular systems. A “top-down” approach (or step-wise design) describes the breaking down, or decomposing, of a system to construct the material of interest or to gain insight into its compositional subunits. Nanoscale materials are carved into shape by physical methods such as lithography. In contrast, a “bottom-up” approach describes the piecing together, or synthesis, of fragments to form the larger molecular system. In the most fundamental sense, nanotechnology refers to the ability to construct molecular assemblies from the bottom up, using current methods and tools to make useful products. The bottom-up approach may take advantage of specific chemical reactions or may involve intermolecular interactions between molecular fragments. Self-assembly, which is one of the important terms used in nanoscience, describes a process in which a collection of disordered building blocks (molecules or nano-objects) come together to form an organized structure. One common example is the crystallization of small ions into a definite lattice structure. Techniques such as x-ray diffraction show that such highly organized structures are comprised of a repetition of smaller “unit cells” of nanoscale dimension. Self-assembled structures may also result from just a few molecules, but the cooperative interplay between these molecules within the structure may impart a specific function or property.

It is also worthwhile to note that self-assembly may lead to discrete or extended entities. A discrete entity is well defined in terms of the number of molecules it contains. A simple example is the dimer that is formed when two acetic acid molecules interact through hydrogen bonding (Figure 1.2). Another good example, which is discussed in detail later, is a surfactant “micelle”—a stable, often spherically shaped aggregate comprised of a few hundred molecules. An extended entity is undefined in at least one dimension. An example is a thin film with a well-defined thickness, say, one molecule thick, but with undefined length and width. A polymer molecule, in which the exact number of monomeric building blocks is



**Figure 1.2** The acetic acid dimer. The dashed lines represent hydrogen bond interactions (discussed in Chapter 2) between the carbonyl oxygen and the hydroxyl hydrogen. The dimer is an example of a discrete entity.

usually unknown, is another example of an extended entity. Nanotechnology exploits both discrete and extended entities for specific functions.

Although scientists use artificial methods of self-assembly to construct functional materials, there are many examples of natural self-assembly processes. The folding of proteins and other biomacromolecules and the formation of lipid bilayer cell membranes are examples of natural self-assembly in biology. Artificial methods may involve covalent building blocks based on chemically robust starting materials or may exploit the shapes, sizes, and intermolecular interactions of these building blocks to direct the formation of nanostructures.

Self-assembly processes can be static or dynamic. Static self-assembly describes the irreversible formation of a stable structure. Examples of static self-assembly include the formation of the double helix of DNA and the folding of polypeptide chains into a protein molecule. Dynamic self-assembly describes a reversible process such as adsorption of molecules onto a surface and oscillating chemical reactions.

## 1.4 SUPRAMOLECULAR SCIENCE

Self-assembly in nanomaterials falls into the broader field of supramolecular science. Supramolecular science refers to the branch of science that focuses on systems composed of a discrete number of molecular subunits (typically of nanoscale dimension). These subunits are sometimes referred to as molecular building blocks. Generally speaking, the spatial organization of the building blocks is influenced by reversible weak interactions such as hydrogen bonds, van der Waals interactions, and electrostatic forces (Chapter 2). Although irreversible interactions such as covalent bonds may also play a vital role, supramolecular chemistry is concerned mainly with non-covalent interactions.

Supramolecular science is important in a host of processes such as protein folding, molecular recognition, self-assembly, and host-guest chemistry. The acetic acid dimer mentioned earlier is a discrete supramolecular entity. The hybridization of single-stranded DNA in solution to the double-stranded form is driven by hydrogen bonds being formed between base pairs—this process results in an extended supramolecular entity.

Supramolecular science and the study of non-covalent interactions touches every scientific discipline from biology (e.g., biological cell structure, protein-protein interactions, drug delivery using nano-vehicles), chemistry (e.g., colloid stability,

micellar nano-reactor synthesis), and physics (e.g., organic photovoltaic systems, holography, optical coatings, data transmission and storage), to engineering (e.g., tertiary oil recovery, large-scale synthesis of nano-wires) and environmental science (e.g., remediation on nano-pores, detection of hazardous materials on nano-films). For example, an extended entity such as a thin film containing an antigen on the surface may be used to detect the presence of a specific antibody. The complexation of the antigen and the antibody, driven by non-covalent interactions, may result in a supramolecular “bilayer.”

In this book, the terms supramolecular chemistry and self-assembly are always used in the context of nanomaterials.

## 1.5 SOURCES OF INFORMATION ON NANOSCIENCE

In this book each chapter ends with a short list of relevant further reading. These books, review articles, and papers have been selected in part due to their clarity, depth, and manageable mathematical rigor. They should be accessible reading material once the student has completed relevant chapters in this text. However, some advanced books, which are typically used by graduate students, technicians, or academic and industrial professionals, are also listed for the ambitious student.

Readers are also encouraged to browse scientific journals for up-to-date information on nanomaterials. The following journals are highly recommended: *Nano Letters*, *Langmuir*, *Journal of Physical Chemistry*, *Biomacromolecules*, *Advanced Materials*, *ACS Nano*, *Applied Materials and Interfaces*, and *Chemistry of Materials* (all published by the American Chemical Society). Other useful journals include *Thin Solid Films*, *Nano Today*, *Nanomedicine* (Elsevier), *Soft Matter*, *Nanoscale* (Royal Society of Chemistry), *Nature Nanotechnology*, and *Nature Materials* (Nature Publishing Group).

The following books are additional recommended starting points for students beginning their education in nanoscience.

- Deffeyes, K. S. and Deffeyes, S. E. *Nanoscale: Visualizing an Invisible World*, 2009, Massachusetts Institute of Technology, Cambridge, MA. This is a beautifully illustrated book containing many examples of nanomaterials. The book describes the local structure of materials at the nanoscale. It is an excellent read for those beginning their studies in nanoscience.
- Ratner, M. A. *Nanotechnology: A Gentle Introduction to the Next Big Idea*, 2002, Prentice Hall, New Jersey. This

book focuses on the technical and business aspects of the field. It provides a wide perspective on the subject, from science and economics to ethics.

- Jones, R. A. *Soft Machines: Nanotechnology and Life*, 2008, Oxford University Press, New York, NY. Although not mathematically rigorous, this book does an excellent job of presenting fundamental physical laws governing nanoscience.
- *Understanding Nanotechnology* from the editors of *Scientific American*, 2002, Warner Books, New York, NY. This is a popular science book that does a good job of describing the technological implications of nanoscience. The mathematical and scientific background is limited in this book, so it is an accessible overview of the field and nanoscience terminology.

## Two

---

# Intermolecular Interactions and Self-Assembly

---

### CHAPTER OVERVIEW

Nanostructures assemble, often spontaneously, from simple molecular building blocks. It is therefore important to begin this chapter with a discussion of the forces between such molecules. The types of intermolecular interactions (for example ion-ion, ion-dipole, dipole-dipole, dipole-induced dipole, London forces, hydrogen bonds, and electrostatic forces) will ultimately determine the degree and type of intermolecular aggregation, as well as the structure of the resulting aggregate. Such interactions are examined in both bulk media and on surfaces. This chapter concludes with some coverage on electronic structure, and how simple quantum mechanical models can be applied to predict some optical properties of nanomaterials. In particular, conjugation in simple organic molecules is used to make important connections between electronic structure, intermolecular interactions, and molecular self-assembly.

### 2.1 INTERMOLECULAR FORCES AND SELF-ASSEMBLY

This section introduces selected fundamental physical ideas relating to the assembly and properties of nanomaterials in order to provide a sufficient background for understanding subsequent chapters. Intermolecular interactions play a central role in surface chemistry and the process of self-assembly, both of which affect the structure and properties of nanomaterials. Such interactions also determine the properties of surfactants, influence adsorption phenomena, and even affect the interaction between molecules and electromagnetic radiation.

Self-assembly is the process during which molecular fragments spontaneously and often reversibly organize themselves into nanomaterials. The organization of these molecular

building blocks is driven by a combination of thermodynamic factors, kinetic factors, and intermolecular interactions. The interactions may be covalent in nature, leading to strong bonds between the molecules and resulting in an irreversibly self-assembled nanostructure. Covalent interactions are mentioned throughout this chapter with specific examples given in Chapter 5. It is important to appreciate that self-assembly can be spontaneous and directed. We begin this chapter by reviewing some important non-covalent intermolecular interactions that govern the formation of self-assembled nanomaterials.

Various forces are responsible for intermolecular interactions. Most of the forces are electrostatic in origin, and we discuss them from a classical perspective, although it should be noted that a quantum mechanical approach to understanding intermolecular forces is perhaps more correct.

Any interaction between two molecules can be thought of as a sum of a variety of different forces. We will discuss many of these forces, including ion-ion forces, ion-dipole forces, dipole-dipole forces, induced dipole forces, dispersion forces, and hydrogen bonds. Depending on the types of molecules interacting, one force or another may predominate.

Scientists often express intermolecular interactions not as forces, but as intermolecular potentials (or the potential energy of interaction). The potential energy ( $V$ ) and force ( $F$ ) between two interacting molecules are related by

$$F(r) = -\frac{dV(r)}{dr} \quad (2.1)$$

where  $r$  is the distance between the two molecules and may be defined differently for different types of interacting molecules. The negative sign on the derivative means that as the potential energy of the interacting molecules increases with increasing  $r$ , the force will tend to move it toward smaller  $r$  to decrease the potential energy.

The van der Waals interaction is the collective term used to describe attractive or repulsive forces, or non-covalent interactions, between molecules. Named after Dutch scientist Johannes Diderick van der Waals, this type of intermolecular interaction generally refers to molecules involving ion-ion, ion-dipole, dipole-dipole forces, and interactions involving induced dipoles (including London dispersion forces). Van der Waals forces play a key role in biology, polymer science, surface science, nanotechnology, and material science. They

govern self-assembly processes, protein-protein interactions, and crystallization processes. These interactions are also found in nature. For example, the ability of geckos to climb smooth surfaces (such as glass) is attributed to van der Waals interactions and likely involves a nanofilm of water trapped between the glass surface and the foot. In fact, research is currently being done in many nanoscience laboratories to mimic this behavior and allow people to scale walls, or to create “gecko tape” that exploits this ability.

Geckos possess the ability to cling to nearly any surface, sometimes even by a single toe, because they have millions of branching hairs called setae on their toe-pads that present enough surface area for van der Waals interactions to have an influence at the macroscopic scale. Recent advances in nanotechnology have yielded reusable “adhesives” that are four times more “adhesive” than a gecko’s foot. These adhesives are comprised of flexible polymers connected by silicon bases to carbon nanotubes, which are cylindrical graphene columns held together by van der Waals interactions. Structure, properties, and uses of carbon nanotubes will be further discussed in Chapter 5.

### 2.1.1 Ion-Ion Interactions

Ion-ion forces are perhaps the most well-known intermolecular forces and are among the strongest intermolecular forces of those we’ll be discussing. Ion-ion forces arise between two ionic (charged) species, such as the force between  $\text{Na}^+$  and  $\text{Cl}^-$  that holds together crystals of common table salt. The potential energy of interaction  $V(r)$  between two charges  $q_1$  and  $q_2$  is often called the Coulombic energy and is given as

$$V(r) = \frac{q_1 q_2}{4\pi \epsilon_0 r_{12}} \quad (2.2)$$

where  $\epsilon_0$  is the permittivity of free space ( $8.854 \times 10^{-12} \text{ m}^{-3} \text{ kg}^{-1} \text{ s}^4 \text{ A}^2$ ) and  $r_{12}$  is the distance between the two ionic species. For atomic or molecular ions,  $q$  is often calculated as  $q = ze$  where  $z$  is the formal charge on the ion and  $e$  is the charge on an electron,  $1.60217 \times 10^{-19} \text{ C}$ .

We can express the Coulombic force using the relationship between potential energy and force (Equation 2.1) as

$$F(r) = -\frac{dV(r)}{dr} = \frac{q_1 q_2}{4\pi \epsilon_0 r_{12}^2} \quad (2.3)$$



Equation 2.3 is obtained by differentiating Equation 2.2, and realizing that  $d/dr(1/r)$  is  $-1/r^2$ . From Equation 2.3, we see that the Coulombic force between two ions changes as  $1/r^2$ . We also see that the force is negative when the two ions are attracted to each other (when  $q_1$  and  $q_2$  have opposite signs) and positive when they repel each other.

### Example 2.1 The Coulombic Energy Between Ions

The ionic radius of  $\text{Na}^+$  is determined to be 95 pm and the ionic radius of  $\text{Cl}^-$  is 181 pm. Calculate the Coulombic energy between two isolated ions of  $\text{Na}^+$  and  $\text{Cl}^-$  if they are in contact as shown in Figure 2.1.

*Solution* If the two ions are “in contact,” then the distance between their centers is the sum of their two ionic radii. So,

$$r_{\text{Na-Cl}} = 95 \text{ pm} + 181 \text{ pm} = 2.76 \times 10^{-10} \text{ m}$$

Each ion has a formal charge of +1 or -1, so

$$q_{\text{Na}^+} = ze = (+1)(1.60217 \times 10^{-19} \text{ C}) = 1.60217 \times 10^{-19} \text{ C}$$

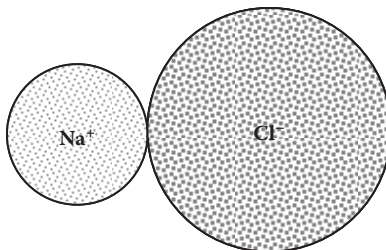
$$q_{\text{Cl}^-} = ze = (-1)(1.60217 \times 10^{-19} \text{ C}) = -1.60217 \times 10^{-19} \text{ C}$$

Then, using Equation 1.2,

$$\begin{aligned} V(r)_{\text{Na-Cl}} &= \frac{q_{\text{Na}^+} q_{\text{Cl}^-}}{4\pi \epsilon_0 r_{\text{Na-Cl}}} = \frac{(1.602 \times 10^{-19} \text{ C})(-1.602 \times 10^{-19} \text{ C})}{4\pi(8.854 \times 10^{-12} \text{ m}^{-3} \text{ kg}^{-1} \text{ s}^4 \text{ A}^2)(2.76 \times 10^{-10} \text{ m})} \\ &= -8.36 \times 10^{-19} \text{ J} \end{aligned}$$

where  $1 \text{ J} = 1 \text{ kg m}^2 \text{ s}^{-2}$  and  $1 \text{ C} = 1 \text{ A s}$ .

One example of ion-ion interactions can be found in the formation of a charged polymer layer, or nanofilm, on a silicon surface. Silicon usually has an oxide layer ( $\text{SiO}_2$ ) of about 1 nm

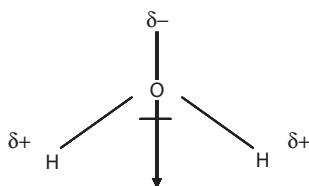


**Figure 2.1** A schematic depiction of  $\text{Na}^+$  and  $\text{Cl}^-$  “in contact” with each other.

in thickness on its surface. In a network of surface  $\text{SiO}_2$  groups, each silicon atom has a tetrahedral molecular geometry, resulting in a layer of oxygen atoms each covalently bonded to a single silicon atom at the surface. At very low pH, these oxygen atoms are protonated, but at neutral to high pH, the oxygen atoms are deprotonated and thus create a layer of negative ionic charge ( $\text{Si-O}^-$ ) along the surface. Polyethylenimine (PEI), a polycation with several amine functional groups, can then be exposed to this negatively charged surface to create a positively charged PEI layer of relatively uniform thickness. A negatively charged ion or polymer can then be exposed to this surface-bound PEI layer to create a secondary layer, resulting in a new negatively charged surface that allows the process to be repeated. This process is called electrostatic self-assembly and will be used to develop several techniques in future chapters.

### 2.1.2 Ion-Dipole Interactions

Many molecules possess permanent dipoles and are classified as polar molecules. Polar molecules do not have a permanent charge, but because of the differing electronegativities of the atoms bound in the molecule, certain regions of the molecule may have a partial positive or a partial negative charge. In certain cases this partial charge can lead to a permanent dipole. For example, a water molecule has a permanent dipole due to its bent geometry. The oxygen atom has a much higher electronegativity than the hydrogen atoms, and so it tends to draw more electrical charge to itself. As a result, the hydrogens have a partial positive charge and the oxygen has a partial negative charge. The net dipole moment passes through the oxygen atom and bisects the hydrogen atoms as shown in Figure 2.2.



**Figure 2.2** The electronegative oxygen pulls electron density away from the hydrogen atoms, leaving them with a partial positive charge  $\delta+$ . Each of the hydrogens has a dipole moment that points toward the partial negative charge of the oxygen, resulting in a net dipole moment that passes through the O and bisects the Hs.

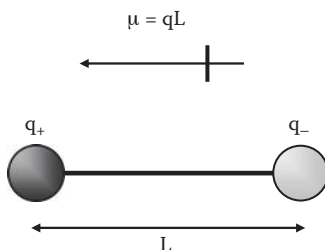
Polar molecules such as  $\text{H}_2\text{O}$  are often characterized by their dipole moment  $\mu$ , which is defined as

$$\mu = qL \quad (2.4)$$

where  $L$  is the distance separating the partial positive and partial negative charges of magnitude  $q$  as shown in Figure 2.3. Dipole moments of some common molecules are shown in Table 2.1. As shown in the table, dipole moments are often given in units of Debye (D), where  $1\text{D} = 3.336 \times 10^{-30} \text{ C m}$ .

When an ion draws near to a molecule with a dipole, there is an electrostatic interaction between the dipole and the ionic species. The potential energy of this interaction is given as

$$V(r, \theta) = \frac{q_{\text{ion}} \mu \cos \theta}{4\pi \epsilon_0 r_{12}^2} \quad (2.5)$$

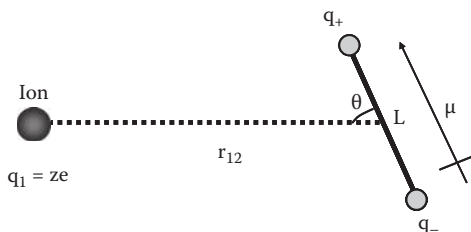


**Figure 2.3** The dipole moment  $\mu$  between two charges  $q_+$  and  $q_-$  separated by a distance  $L$  is calculated as  $\mu = qL$ .

**Table 2.1** Dipole Moments of Some Common Molecules

Molecule	Dipole Moment (D)
$\text{H}_2\text{O}$	1.85
$\text{CH}_3\text{OH}$	1.7
$\text{NH}_3$	1.47
$\text{CH}_3\text{Cl}$	1.9
$\text{CHCl}_3$	1.04
$\text{CH}_3\text{COOH}$	1.7
$\text{NaCl}$	9.00
$\text{HCl}$	1.11
$\text{C}_6\text{H}_5\text{CN}$ (benzonitrile)	4.18

Source: *CRC Handbook of Chemistry and Physics*, 88th ed., Web Version. (2008) pp. 9-47–9-55. With permission.



**Figure 2.4** A schematic depiction of the variables involved in an ion-dipole interaction.  $L$  is the distance between the centers of the two partial charges of the dipole.  $r_{12}$  is the distance between the center of the ion and the midpoint of  $L$ .  $\theta$  is the angle between  $L$  and  $r_{12}$ .  $q_1$  is the charge on the ion, calculated as  $ze$ .  $\mu$  is the dipole moment.

where  $\mu$  is the dipole moment,  $q_{\text{ion}}$  is the charge on the ion,  $r_{12}$  is the distance between the ion and the center of the dipole moment, and  $\theta$  is the angle between  $L$  and  $r_{12}$ . A schematic of the geometries involved in a typical ion-dipole interaction is shown in Figure 2.4.

From Equation 2.5 we see that the potential energy of an ion-dipole interaction is angle dependent, which makes sense intuitively. For example, consider the interaction between a cation (a positively charged ion) and a dipole. The cation attracts the negative region of the dipole, but repels the positive region. If the negative region is oriented toward the cation and the positive region is pointing away from it ( $\theta = \pi$ ), we expect the magnitude of the potential energy to be maximized. Likewise, if the positive region is pointing toward the cation and the negative region is pointing away ( $\theta = 0$ ), the magnitude of the potential energy is maximized (but it is repulsive rather than attractive). If, however, the dipole is oriented perpendicularly with respect to the cation ( $\theta = \pi/2$ ), the potential energy is zero because the attractive interaction between the cation and the negative region is balanced out by the repulsive interaction with the positive region. Therefore, the angle dependence of the ion-dipole interaction is as expected.

It should be noted that some molecules (such as benzene or  $\text{CO}_2$ ) may have partial charge separations, but do not have a net dipole moment. For example, in the case of  $\text{CO}_2$ , the dipole moment from the carbon toward one oxygen cancels out the dipole moment from the other. However, in these types of molecules, a *multipole* may exist and also results in an intermolecular force between the molecule and an ionic species. Different equations than those above must be used to calculate

the interaction between molecules with electrical multipoles and ionic species.

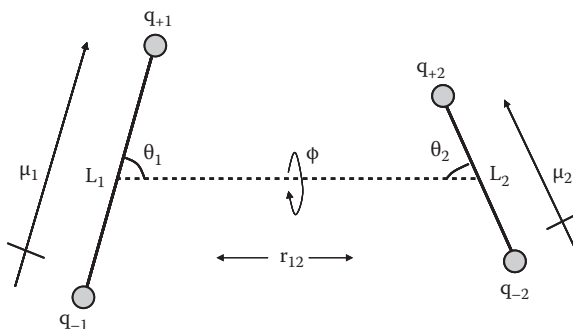
### 2.1.3 Dipole-Dipole Interactions

Molecules with permanent dipoles may also interact with each other through electrostatic means. As might be expected, the strength of the interaction is also angle dependent. This type of interaction is analogous to the magnetic attraction between two bar magnets—the attraction between the two magnets depends on the angle of rotation of each magnet relative to the other. A schematic representing the interaction between two molecular dipoles is depicted in Figure 2.5. The potential energy for such an interaction between two dipole moments  $\mu_1$  and  $\mu_2$  can be calculated as

$$V(r, \theta_1, \theta_2, \phi) = -\frac{\mu_1 \mu_2}{4\pi \epsilon_0 r_{12}^3} (2 \cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2 \cos \phi) \quad (2.6)$$

where  $\theta_1$ ,  $\theta_2$ , and  $\phi$  are defined in Figure 2.5.

It might be tempting to think that the attraction between two dipoles is always maximized when the two dipoles are in line “head to tail” with the partial positive charge of one dipole pointing directly toward the partial negative charge of the other. However, depending on the lengths  $L$  of the interacting dipoles, the most attractive interaction may often be when the two dipoles are anti-parallel to each other, with the positive region of one dipole directly adjacent to the negative



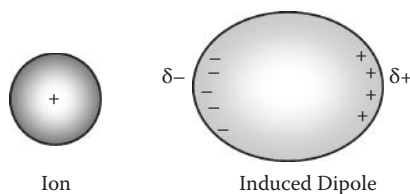
**Figure 2.5** A schematic depiction of the variables involved in a dipole-dipole interaction.  $L$  is the distance between the centers of the two partial charges of either dipole 1 or 2.  $r_{12}$  is the distance between the midpoints of  $L_1$  and  $L_2$ .  $\theta_1$  and  $\theta_2$  are the angles between  $L_1$  and  $r_{12}$  or  $L_2$  and  $r_{12}$ , respectively.  $\phi$  is the angle of rotation between dipoles 1 and 2.  $q_+$  and  $q_-$  are the partial positive or negative charges in each dipole.  $\mu$  is the dipole moment.

region of the other. This anti-parallel orientation allows the molecules to draw closer together, reducing the value of  $r_{12}$  and thereby maximizing the attractive interaction energy.

In the purification technique of column chromatography, solid polar molecules such as silica or alumina are placed in a vertical glass column and are referred to as the stationary phase. The liquid mobile phase of solution to be purified is then flowed through the column. Because dipole-dipole interactions between polar molecules in the mobile phase and the molecules of the stationary phase slow the polar molecules' descent, compounds in the solution flow out the bottom of the column, or elute, in order of increasing polarity and can thus be separated. This same technique works for ions in solution as well because they are slowed by ion-dipole interactions. Polarity of the solvent dictates the rate of movement of compounds through the column. If a solvent is too polar, the stationary phase attracts the solvent rather than the solutes and no separation occurs. However, if a solvent is not polar enough, some more polar solutes may not make it all the way through the column.

### 2.1.4 Interactions Involving Induced Dipoles

When an ion approaches a non-polar molecule, the electrons of the non-polar molecule may feel the effect of the electric field produced by the ion. As a result, the electron cloud surrounding the non-polar molecule may become distorted. For example, when a cation approaches a non-polar molecule, the electron cloud of the non-polar molecule is pulled slightly toward the cation. The result of this distortion of the electron cloud is an effective charge separation in the non-polar molecule and is called an induced dipole. A schematic of an ion-induced dipole interaction is shown in Figure 2.6. Polar molecules, as well as ions, are also capable of inducing dipoles in non-polar molecules.



**Figure 2.6** As a cation approaches a polarizable atom or molecule, its electric field produces a distortion of the electron cloud surrounding the polarizable atom or molecule. The proximity of the cation to the molecule results in an effective charge separation and produces an induced dipole in the molecule.

The extent to which the electron cloud of a molecule becomes distorted in the presence of an ion or polar molecule is termed its polarizability. Polarizability  $\alpha$  is mathematically defined in terms of the strength of the dipole induced in a molecule due to an electric field of strength  $E$ ,

$$\mu_{\text{induced}} = \alpha E \quad (2.7)$$

Molecules with high polarizabilities have a larger induced dipole moment in the presence of an electric field than those with low polarizabilities. Typical polarizabilities of various atoms and molecules are shown in Table 2.2. Note that the units of polarizability given in the table are  $10^{-24} \text{ cm}^3$  divided by  $4\pi\epsilon_0$ .

The presence of an induced dipole moment in a non-polar molecule means that a potential energy of interaction exists between the non-polar molecule and the ion or polar molecule that is inducing the dipole. Using Coulomb's law and Equation 2.3, we can calculate the electric field produced by an ion as a function of the distance  $r$  from the center of the ion

$$E(r) = \frac{F}{q} = \frac{q}{4\pi\epsilon_0 r^2} \quad (2.8)$$

The interaction energy between an ion and an induced dipole can then be given as

$$V(r) = \frac{-q_1\alpha}{2(4\pi\epsilon_0)^2 r_{12}^4} \quad (2.9)$$

**Table 2.2** Polarizabilities of Several Atoms and Molecules

Molecule	Polarizability ( $10^{-24} \text{ cm}^3/4\pi\epsilon_0$ )
H <sub>2</sub> O	1.45
D <sub>2</sub> O	1.26
NH <sub>3</sub>	2.2
He	0.20
H <sub>2</sub>	0.8
CH <sub>4</sub>	2.59
Au	5.8
Si	5.38
C <sub>6</sub> H <sub>6</sub> (benzene)	10.2

Source: *CRC Handbook of Chemistry and Physics*, 88th ed., Web Version. (2008) pp. 10-193–10-202. With permission.

### Example 2.2 Perturbation of Electron Clouds Due to Ions

The center of a sodium ion ( $\text{Na}^+$ ) is located 0.35 nm from the center of a gold atom. If the atomic radius of gold is 144 pm, by what percentage of its atomic radius would the electron cloud of the gold atom be shifted due to the presence of the sodium ion?

*Solution* The electric field induced by  $\text{Na}^+$  at a distance 0.35 nm from its center is calculated using Equation 2.8.

$$E(r) = \frac{q_{\text{Na}^+}}{4\pi\epsilon_0 r^2} = \frac{(1.602 \times 10^{-19} \text{ C})}{4\pi(8.854 \times 10^{-12} \text{ m}^{-3} \text{ kg}^{-1} \text{ s}^4 \text{ A}^2)(0.35 \times 10^{-9} \text{ m})^2}$$

$$= 1.18 \times 10^{10} \frac{\text{J}}{\text{Cm}}$$

Then, using Table 1.1, the dipole moment induced in the gold atom is

$$\mu = \alpha E = 4\pi\epsilon_0 (5.8 \times 10^{-30} \text{ m}^3) \left( 1.18 \times 10^{10} \frac{\text{J}}{\text{Cm}} \right) = 7.61 \times 10^{-30} \text{ Cm}$$

Now, from Equation 2.4, we know that  $\mu = qL$ , and so per unit charge ( $e$ ) we can calculate  $L$

$$L = \frac{\mu}{q} = \frac{7.61 \times 10^{-30} \text{ Cm}}{1.602 \times 10^{-19} \text{ C}} = 4.75 \times 10^{-11} \text{ m} = 47.5 \text{ pm}$$

Then the electron cloud of Au is shifted by 47.5 pm/144 pm = 33% of its atomic radius.

In an analogous manner, the electric field strength at a given point in space produced by a polar molecule with dipole moment  $\mu$  is a function of the orientation of the dipole moment with respect to that point in space and is calculated as

$$E(r, \theta) = \frac{\mu(3\cos^2\theta + 1)^{1/2}}{4\pi\epsilon_0 r^3} \quad (2.10)$$

The potential energy of interaction between a polar molecule with permanent dipole  $\mu_1$  and an induced dipole is therefore

$$V(r, \theta) = \frac{-\mu_1^2 \alpha (3\cos^2\theta + 1)}{2(4\pi\epsilon_0)^2 r_{12}^6} \quad (2.11)$$



where  $\theta$  is the angle between the dipole moment of the polar molecule and the line connecting the midpoint of the polar molecule with the center of the induced dipole.

Finally, we note that the interaction between an ion or a polar molecule and an induced dipole is always attractive. It is inherently so because the electric field produced by the ion or polar molecule always induces a dipole in the polarizable molecule that is oriented such that it is attracted toward the species inducing the dipole.

### 2.1.5 Dispersion Forces

Aside from the forces that are essentially electrostatic in nature described in the previous sections, there also exists a force between all molecules, regardless of charge or polarity, that results from the quantum mechanical correlation between electrons in neighboring molecules. This force is called the dispersion or London force. Although dispersion forces are inherently quantum mechanical in nature and a rigorous description of their origin is beyond the scope of this book, we can gain an intuitive grasp of dispersion forces by considering in a somewhat classical manner their contribution to the interaction between two neutral, non-polar molecules.

Even though a neutral, non-polar molecule has no permanent dipole moment, at any given moment the distribution of its electrons may be asymmetrical, resulting in an instantaneous or momentary dipole moment. This instantaneous dipole moment creates an electric field that perturbs the electrons of a neighboring molecule, producing in it an induced dipole moment and resulting in an attractive force between the two molecules.

In order to calculate the dispersion force between two molecules, a quantum mechanical calculation must be performed, the accuracy of which generally depends on the level of theory used. One of the earliest calculations was performed by London in the 1930s using quantum mechanical perturbation theory. His calculation produces reasonably accurate results, and although more precise calculations have been performed in more recent years, London's equation is less complex and therefore more suitable for our purposes.

According to London's results, the approximate potential energy of interaction due to dispersion between two molecules can be calculated in terms of their electronic polarizabilities  $\alpha$  and their ionization potentials  $I$ . For two identical molecules (or atoms), the result is

$$V(r) = \frac{-3}{4(4\pi\epsilon_0)^2} \frac{\alpha^2 I}{r_{12}^6} = \frac{-C_{\text{dispersion}}}{r_{12}^6} \quad (2.12)$$

and for two different molecules

$$V(r) = \frac{-3}{2(4\pi\epsilon_0)^2} \frac{\alpha_1\alpha_2}{r_{12}^6} \frac{I_1 I_2}{(I_1 + I_2)} \quad (2.13)$$

As with the interaction energy for dipole-induced dipole interactions in Section 2.1.4 (Equation 2.11), we see that the potential energy of interaction for a dispersion interaction according to London's equations goes as  $1/r^6$  and is always attractive between any two molecules.

The dispersion interaction plays an important role in the liquid and solid phases of many materials and is the main contributor to cohesion. However, it turns out that the strength of the dispersion interaction does not vary much between different types of molecules (i.e., the interaction between any two given molecules is of similar strength). Therefore, the electrostatic interactions described in the earlier sections, and not dispersion interactions, are generally responsible for such behaviors as phase separation and self-assembly in condensed phases, behaviors that are of utmost importance in the development and study of nanomaterials.

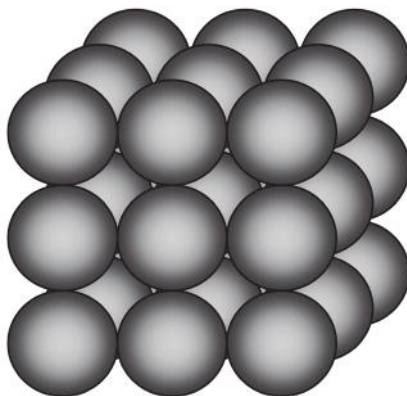
### 2.1.6 Overlap Repulsion

In our discussion of the different types of intermolecular potentials in the previous sections, we ignored the fact that atoms and molecules occupy some finite space. For example, if we examine the equation for Coulombic force (Equation 2.3) by itself, we would be led to conclude that two oppositely charged ions are drawn toward each other with increasing force until they occupy the same point in space. Obviously, this does not occur with atoms and molecules in nature. To account for the finite size of atoms and molecules, we then include another contributor to the interaction potential energy between two atoms or molecules called overlap repulsion. Overlap repulsion is the interaction that accounts for the fact that two atoms or molecules cannot occupy the same point in space.

What, then, is the size of an atom or a molecule? This is not a trivial question. From the results of quantum mechanics, we realize that the electron "clouds" of atoms and molecules do not have definite boundaries. Determining where an atom

“ends” is therefore somewhat tricky. As a result, radii of atoms are often experimentally defined and, depending on the type of measurement made (and consequently the property measured), a different result might be obtained. For example, one way to measure the radii of atoms would be to assume that in solids they act as tiny, hard spheres that have packed closely together (see Figure 2.7). Using x-ray or neutron diffraction methods, one could then observe how closely the atoms pack together in a crystal and thereby deduce the atomic radius. The results of this method yield a type of atomic radius called the *hard sphere radius* or *van der Waals packing radius*. Other methods also exist, such as measuring the distance between two atoms in a covalent bond (rather than the distance in a crystal). This method yields the *covalent bond radius*. The atomic radius calculated depends on the method used. In certain cases, the results obtained from these different methods might vary by as much as 30%. The type of measurement one chooses to use generally depends on the type of system being studied.

After using the most suitable method to determine an atomic radius, one can then calculate the overlap repulsion between two atoms. A variety of models with increasing sophistication and complexity are used to calculate the repulsive potential energy between two atoms due to overlap. Perhaps the simplest model is to characterize the atom as a “hard sphere” with a definite boundary (i.e., the repulsive force between two atoms would be infinite at any distance smaller than the atomic radius). This hard sphere model between two atoms at



**Figure 2.7** Atoms in crystal lattices can often be modeled as tiny, hard spheres in order to calculate their atomic radius. X-ray or neutron diffraction methods can then be used to experimentally determine the atomic radius.

a distance  $r$  from each other can be represented mathematically as

$$V(r) = \left( \frac{\sigma}{r} \right)^\infty \quad (2.14)$$

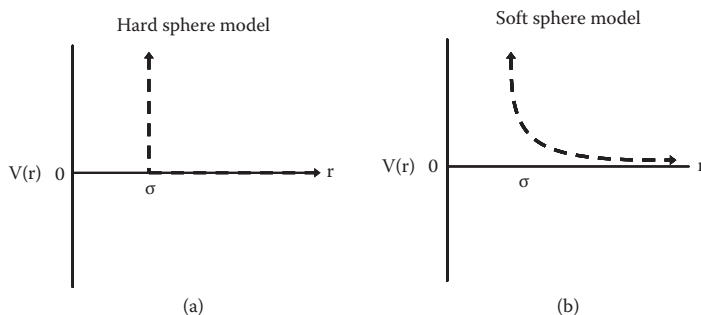
where  $\sigma$  is the atomic or molecular diameter (i.e., two times the atomic radius). As expected, when  $r > \sigma$ , then  $V(r)$  is essentially zero and when  $r < \sigma$ ,  $V(r)$  is infinitely large. A graph of  $V(r)$  versus  $r$  for the hard sphere model is shown in Figure 2.8(a).

A more realistic model is the soft sphere model, which assumes that atoms are “compressible” to some degree and do not have completely rigid boundaries. One mathematical representation for the soft sphere model can be given as a power law as

$$V(r) = \left( \frac{\sigma}{r} \right)^n \quad (2.15)$$

where  $n$  is usually an integer between 9 and 16 and  $\sigma$  is defined as before. In this model,  $V(r)$  quickly becomes quite small when  $r$  is much bigger than  $\sigma$ . Conversely, the overlap repulsion becomes large rather quickly when the distance between the two atomic centers is less than  $\sigma$ . Figure 2.8(b) depicts a graph of  $V(r)$  versus  $r$  for the soft sphere power law model.

Thus far in our discussion of overlap repulsion, we have operated under the assumption that the atomic or molecular



**Figure 2.8** (a) The hard sphere model of overlap repulsion.  $r$  is the intermolecular distance and  $\sigma$  is the molecular diameter. (b) The soft sphere model of overlap repulsion. Note that  $r$  can assume some values slightly smaller than  $\sigma$  without  $V(r)$  becoming infinitely large, as is the case with the hard sphere model.

geometry is essentially spherical. While this assumption is relatively sound for atoms and some small molecules ( $\text{CH}_4$ , for instance, can be modeled as nearly spherical), most molecules possess other geometries. The concepts of overlap repulsion we have developed still apply to these species, but different calculations of their interaction energy are required to account for their differing geometries. However, such methods of calculation are beyond the scope of this text.

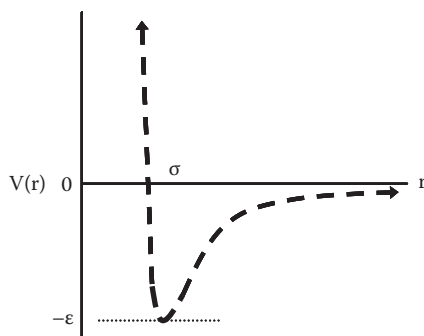
### 2.1.7 Total Intermolecular Potentials

The previous sections provided a fundamental description of the main forces involved in van der Waals intermolecular interactions. Ultimately, the total interaction potential energy between two molecules is the sum of all the different interactions that we have discussed (as well as a few more complex interactions).

In a very basic treatment of the interactions between two atoms or molecules, the total intermolecular potential is often modeled by the Lennard-Jones potential, which is the sum of a soft sphere repulsion term and an attractive term that goes as  $1/r^6$  (analogous to the London dispersion attractive interaction). The Lennard-Jones potential is given as

$$V(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right] \quad (2.16)$$

where  $-\epsilon$  is the minimum energy and  $\sigma$  is a constant parameter (not the molecular diameter). A graph of the Lennard-Jones potential is shown in Figure 2.9.



**Figure 2.9** The Lennard-Jones total intermolecular potential curve.  $-\epsilon$  is the minimum energy.  $r$  is the intermolecular distance.

Although the Lennard-Jones potential is a relatively primitive model of the total intermolecular potential between two molecules, it provides us with a qualitatively useful picture of a common interaction between two molecules. Starting from the far right-hand side of Figure 2.9, we see that the potential energy decreases as the distance between the molecules grows smaller, until the energy reaches a minimum value of  $-\epsilon$ . If  $r$  decreases beyond this minimum energy value, then the potential energy quickly increases (i.e., the force between the two molecules becomes strongly repulsive) due to the overlap repulsion term.

A more complete total intermolecular potential would be the sum of all interaction potential energies. Using only the interactions we have discussed so far, the complete total intermolecular potential would look like that shown in Equation 2.17. However, it must be realized that simple systems will not exhibit all of these kinds of interactions, so that some terms in Equation 2.17 will be zero.

$$\begin{aligned}
 V(r)_{\text{total}} = & V(r)_{\text{ion-ion}} + V(r)_{\text{dipole-dipole}} + V(r)_{\text{ion-dipole}} \\
 & + V(r)_{\text{ion-induced dipole}} + V(r)_{\text{dipole-induced dipole}} \\
 & + V(r)_{\text{dispersion}} + V(r)_{\text{overlap}}
 \end{aligned} \tag{2.17}$$

If the sum of attractive interaction terms is greater than the repulsive interaction terms, then the two molecules are drawn together until the repulsive interactions eventually overwhelm the attractive interactions (remember that the overlap repulsion quickly becomes prohibitively large at distances smaller than the atomic or molecular radii).

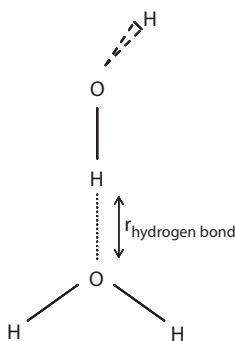
In conclusion and as a cautionary addendum, note that many of the models of intermolecular forces discussed in this and previous sections are for the most part rudimentary and hence not fully useful if one wishes to make exhaustive calculations of the potential energies between atoms or molecules. However, the interactions we have discussed are qualitatively very useful for the purposes of this text and provide the conceptual tools to understand the intermolecular forces at play in the realm of nanomaterials.

### 2.1.8 Hydrogen Bonds

The hydrogen bond is a special type of intermolecular interaction of great importance in colloidal systems and nanomaterials in general. The hydrogen bond is essentially

electrostatic in origin and so is a subset of the dipole interactions already discussed. However, it is of particular importance and strength, and so has acquired a special classification.

Hydrogen bonds occur between molecules that have hydrogen covalently bonded to a strongly electronegative atom such as N, O, or F. In such cases, the electron density surrounding the hydrogen atom is mostly drawn toward the more electronegative atom, leaving the hydrogen atom “exposed” with a strong partial positive charge through which it may form a strong dipole-dipole interaction with the electronegative element on an adjacent molecule. Perhaps more importantly, the reduced electron density around the hydrogen atom means that the neighboring molecule with which it is hydrogen bonding can draw much closer than it could otherwise (within  $\sim 1.5$  to  $2.0$  Å). In other words, the overlap repulsion is minimized because the electron “cloud” surrounding the hydrogen atom has been reduced in size. Because the neighboring molecule can draw much closer, the magnitude of the attractive interaction energy is much greater than it would normally be (remember that in our equations for electrostatic interaction energies,  $r_{12}$  was in the denominator and usually raised to some power; so if  $r_{12}$  is smaller, then the magnitude of the interaction potential energy is much greater). Figure 2.10 depicts a hydrogen bond between two water molecules.



**Figure 2.10** The hydrogen bond between two water molecules. The electronegative oxygen atom pulls much of the electron density surrounding the hydrogen atom to itself, giving the oxygen a large partial-negative charge and leaving the hydrogen atom with a partial-positive charge and very little electron density. The oxygen atom of a neighboring water molecule can therefore approach much closer to the hydrogen atom than would normally be possible for a dipole-dipole interaction.

### 2.1.9 The Hydrophobic Effect

The van der Waals interactions discussed so far are responsible for many of the physical properties (e.g., solubility) of organic molecules. For example, in methanol ( $\text{CH}_3\text{OH}$ ), the hydrocarbon portion of the molecule is relatively small and the polar hydroxyl group is largely responsible for the weak intermolecular van der Waals interactions. However, as the length of the hydrocarbon moiety increases [e.g., as in decanol,  $\text{CH}_3(\text{CH}_2)_9\text{OH}$ ] the non-polar hydrocarbon portion of the molecule dominates the interactions and defines the solubility. Hydrocarbon chains are essentially oil and have little or no tendency to interact with water. If the hydrocarbon chain is long enough the molecules may drop out of solution (precipitate) and interact with themselves instead of the water molecules. This brings us to a discussion of the hydrophobic effect, which describes the tendency of non-polar molecules to form aggregates of like molecules in water.

Experience shows us that at the macroscopic level oil and water do not mix, but rather form separate layers. We also know that water forms beads, or distinct droplets, on oil-like surfaces such as the surface of a leaf. Often when organic molecules containing large non-polar hydrocarbon moieties are placed in water they spontaneously self-organize themselves into larger aggregates because this arrangement minimizes overall contact between the hydrocarbon part of the molecule and the water molecules of the solvent. The origin of this phase separation is primarily to maximize the water-water intermolecular interactions, because the dipole-dipole interactions are stronger than the interactions between non-polar molecules. In other words, the enthalpic force is in the water, not the non-polar phase, meaning that this type of phase separation is enthalpically driven.

In order to complete our understanding of the hydrophobic effect, we need to say something about entropy. In addition to the enthalpic contribution discussed above, entropy may play an important role, especially in more complex systems. From general chemistry, entropy is described as a thermodynamic property related to the number of energy levels among which the energy of a system is spread. Basically this is saying that a collection of water molecules confined and “ordered” within a small region of space will have smaller entropy compared to the same molecules “disordered” and occupying a larger volume. Entropy changes are positive for a spontaneous process, which means that it is favorable for a water molecule to



be “released” from a small space and enter a larger volume, especially if such a process leads to a loss of order.

Consider an organic molecule with a large hydrocarbon chain [e.g.,  $\text{CH}_3(\text{CH}_2)_9\text{OH}$ ] placed in water. It turns out that water molecules, in avoiding contact with the hydrocarbon chain, form a cage around the chain. This cage contains ordered water molecules essentially immobilized around the organic molecule. When two such “hydrated” molecules approach each other and make contact, this cage is disrupted and the confined water molecules are liberated into the bulk solution. This process is accompanied by an increase in entropy of the water molecules. In a sense the van der Waals attractions between the two hydrocarbon chains is driven in part by the large increase in entropy due to the disruption of the cage. If this process occurs among many organic molecules it will lead to aggregation, typically resulting in nanoscale entities dispersed within the aqueous solution.

The hydrophobic effect is paramount in many self-assembly processes, including some biological processes such as the formation of the cell membrane. Reference to the effect is made throughout the text and, in particular, Section 3.3 contains a discussion of the hydrophobic effect in the context of surfactant chemistry.

## 2.2 ELECTROSTATIC FORCES BETWEEN SURFACES: THE ELECTRICAL DOUBLE LAYER

Surface chemistry plays a vital role in the self-assembly of nanomaterials. The forces discussed so far (van der Waals interactions, hydrogen bonds, hydrophobic interactions, etc.) may exist between a planar surface and a molecule some distance away. The strength and nature of these interactions will determine the extent to which molecules adsorb to the surface and perhaps initiate the growth of a nanomaterial. Furthermore, surface forces play an important role in catalysis, where a surface-bound molecule may be immobilized on the surface in an optimal geometry for a reaction to ensue. This section focuses on electrostatic interactions at surfaces. A more thorough treatment of the subject can be found in Israelachvili’s classic book *Intermolecular and Surface Forces*.

### 2.2.1 The Electrical Double Layer

Our discussion of electrostatic forces would not be complete without a brief overview of the electrostatic double layer and its role in the interactions between surfaces in liquids at the

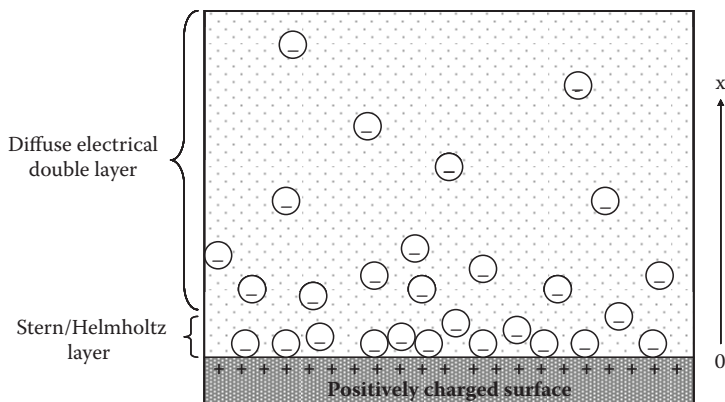
nanoscale. The electrostatic double layer is the term given to the diffuse layer of counterions in a solution that are associated with a charged surface. As shown below, the electrostatic double layer plays an important role in determining the forces that operate between charged surfaces in liquids.

When a surface comes into contact with a liquid, it may become charged by adsorbing ions from solution or releasing ions into solution. Many surfaces, for instance, are pH labile, and at either high or low pH might become positively or negatively charged. For example, a surface containing primary amine groups becomes positively charged at  $\text{pH} < 10$  as the amine groups acquire an extra proton. Another common example of a surface-charging mechanism is the binding of  $\text{Ca}^{2+}$  ions by the zwitterionic head groups of many phospholipid bilayers, resulting in a positively charged surface.

In solution, we would expect the charges on a surface to be balanced by the appropriate counterions that have been released from the surface itself or drawn in from the surrounding solution. Indeed, this is the case and the result is the formation of two regions of counterion charge to neutralize the surface charge. The first region is a compact layer of counterions that is closely bound to the charged surface. This compact region of bound counterions is called the *Stern* or *Helmholtz* layer. It should be noted that the counterions in the Stern layer are not necessarily irreversibly bound to the surface and can often be exchanged with those in the surrounding solution. The second region is a more diffuse and extensive layer of counterions that is in rapid equilibrium with the surrounding solution. This region is referred to as the *electrical double layer* or the *diffuse electrical double layer* and is the focus of our discussion. These regions are shown schematically in Figure 2.11.

A common example of the electrical double layer can be found in the colloid milk. As a mixture of primarily non-polar butterfat droplets in water, milk particles would seem to be expected to aggregate and coagulate into butter due to hydrophobic interactions. However, trace amounts of the highly polar phosphoprotein casein at the water-milk interface result in an electrical double layer forming around each milk particle. This double layer creates enough repulsion to overcome the hydrophobic particle's tendency to aggregate. Inks, paints, and blood provide further examples of heterogeneous liquid mixtures that are stabilized by electrical double layers.

The presence of the electrical double layer is a direct consequence of the tug-of-war between the energy and entropy



**Figure 2.11** The Stern/Helmholtz layer and the diffuse electrical double layer. Ions within the Stern/Helmholtz layer are bound to the surface, although generally not rigidly.

of the charged-surface/bulk solution system. The electrostatic energy of the system is minimized when the charge separation is at a minimum—i.e., when the counterions in solution become closely associated with the charged surface to the point of neutralization. The entropy, on the other hand, is maximized when the counterions are able to move freely through the entire volume of the bulk solution. Based on energy considerations alone, we would not expect an electrical double layer to exist—the surface charges would be completely neutralized by counterions closely bound to the surface. Entropic considerations, however, demand some sort of give and take. The resulting compromise between energy minimization and entropy maximization of the system produces the diffuse double layer with an equilibrium concentration of counterions that gradually decreases with distance away from the charged surface until it reaches a value equivalent to that of the bulk solution.

The actual distribution of the counterions at equilibrium can be calculated by the Poisson-Boltzmann equation

$$\frac{d^2\Psi}{dx^2} = -\left(\frac{e}{\epsilon\epsilon_0}\right) \sum_i z_i \rho_{io} \exp(-z_i e \Psi(x)/kT) \quad (2.18)$$

where  $z_i$  is the valency of the  $i$ -th electrolyte (i.e., +1 for  $\text{Na}^{1+}$ ),  $e$  is the standard unit of charge,  $k$  is Boltzmann's constant, and  $T$  is temperature.  $\Psi(x)$  is the electrostatic potential at a distance  $x$  away from the surface. The zero of the potential can

be arbitrarily assigned and  $\rho_{io}$  is the number density of the  $i$ th electrolyte at that same distance  $x$  (often chosen to be the limit as  $x$  approaches  $\infty$ , which is equivalent to the bulk phase).

### 2.2.2 The Debye Length

For most situations, the solutions to the Poisson-Boltzmann equation are rather complicated and should be obtained numerically by a computer. However, in the limit of a very small electrostatic potential such that  $ze\psi(x)/kT \ll 1$ , then the Poisson-Boltzmann equation reduces to

$$\frac{d^2\psi}{dx^2} = \left( \frac{e}{\epsilon \epsilon_0} \right) \sum_i z_i \rho_{io} (z_i e \psi(x)/kT) \quad (2.19)$$

$$\frac{d^2\psi}{dx^2} = \kappa^2 \psi(x) \quad (2.20)$$

where

$$\kappa = \left( \frac{\sum_i \rho_{io} z_i^2 e^2}{\epsilon \epsilon_0 kT} \right)^{1/2} \quad (2.21)$$

and has units of  $m^{-1}$ . In this case,  $\rho_{io}$  is defined as the number density of the  $i$ th electrolyte in the bulk solution.

The second-order differential equation in 2.20 is called the Debye-Hückel equation and has a well-known solution of

$$\psi(x) = \Psi_0 e^{-\kappa x} \quad (2.22)$$

where  $\Psi_0$  is the potential at the charged surface.

From Equation 2.22, we see that the characteristic decay length of the electrostatic potential for the Debye-Hückel model is  $1/\kappa$ . This length is often called the *Debye length* or the *Debye screening length* and can be used as a rough approximation for the “thickness” of the electrical double layer. If a charge is within the Debye length, it “feels” the effect of the charged surface, and if it is too far outside the Debye length, it effectively is screened from the charged surface by the intervening cloud of counterions.

From Equation 2.21 we also see that the Debye length is independent of the properties of the surface itself—that is to say for a given liquid at a certain temperature it depends only on the concentrations and valencies of ions in the bulk solution, not on the surface charge or surface potential. For example, for an aqueous 100-mM NaCl solution at 25°C, the Debye length is 0.96 nm, independent of the charge density or potential of the surface itself.

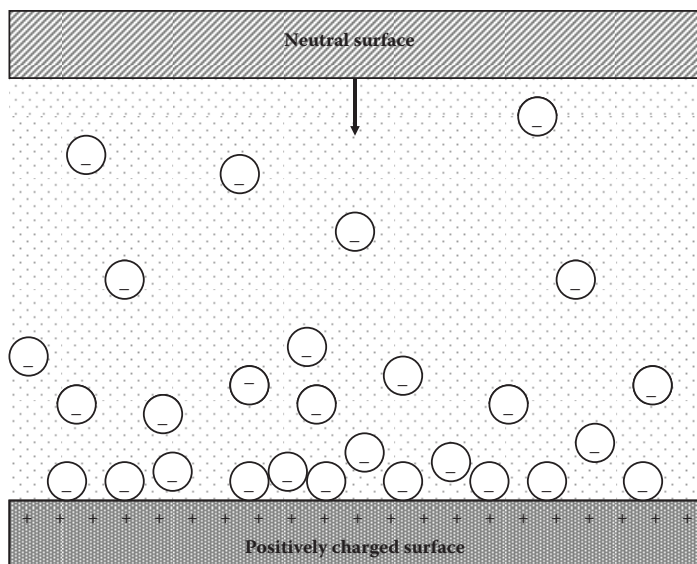
### 2.2.3 Interactions Between Charged Surfaces in a Liquid

In general, we can solve for the pressure due to the presence of ions at a position  $x$  between two charged surfaces as

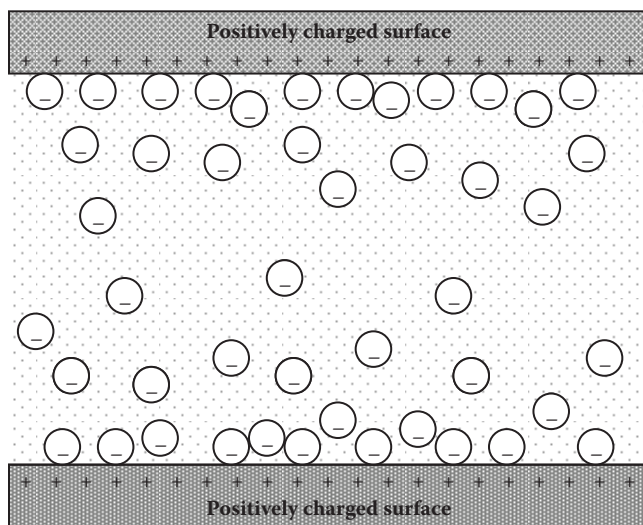
$$P(x) = kT \sum_i \rho_i(x) \quad (2.23)$$

where  $\rho_i(x)$  is the number density of the  $i$ th electrolyte at  $x$  (measured in molecules per cubic meter). Since the distribution of ions at any given point must be calculated using the Poisson-Boltzmann equation, then solving the Poisson-Boltzmann equation for the system must precede a calculation of the pressure between two surfaces. However, as mentioned previously, the Poisson-Boltzmann equation is rather complicated to solve for most systems of practical interest. It is beyond the scope of this text to discuss the solutions to this equation, but students should be aware of this general approach for calculating the pressure between two surfaces. We limit our discussion to qualitative descriptions of the forces operating between two interacting surfaces, descriptions resulting from the application of the approach described previously.

In the most elementary example, suppose we have a flat, neutral surface approaching a flat, charged surface in a parallel orientation, such as that shown in Figure 2.12. Before the approach of the neutral surface, the charged surface has associated with it a diffuse electrical double layer extending out into solution. As the neutral surface approaches, however, the counterions in the double layer must become confined to a smaller and smaller volume, resulting in a decrease in entropy of the system. The approach of the neutral surface likely causes some of the counterions to bind to the charged surface, resulting in a slight decrease in energy. However, this favorable decrease in energy is offset by the much larger decrease in the entropy of the system. For this reason, the interaction



**Figure 2.12** The interaction between a neutral and charged surface is repulsive.



**Figure 2.13** The interaction between two surfaces of like charge is repulsive.

between a neutral and charged surface in a liquid must always be repulsive.

Now let's consider the interaction between two charged surfaces of like charge, as shown in Figure 2.13. As the surfaces approach each other, their electrical double layers begin

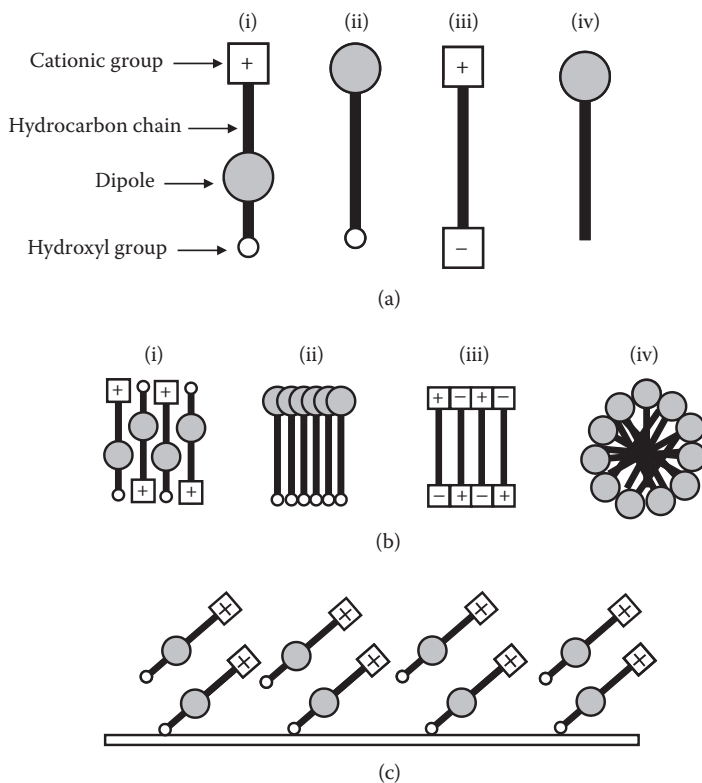
to overlap, resulting in an effective decrease in the entropy of the system and making the interaction unfavorable. Hence, the interaction between two surfaces of like charge is always repulsive. By comparison, this repulsive interaction begins to occur when the two surfaces are at a greater distance than if one of the surfaces was neutral.

Finally, let's consider the case of the interaction between two surfaces of opposite charge. As we would expect, electrostatic attraction between the surfaces dominates at long range. One surface can, in effect, operate as the "counterion" for the apposing surface. As the surfaces approach, counterions are released into solution and expelled from the gap between the two surfaces, resulting in an increase in entropy of the system. If the charge densities of each surface are equivalent, the electrostatic attraction continues up until contact between the surfaces. If the charge densities are not equal, some counterions must remain in the gap. Gradually they become more concentrated as the surfaces approach each other, and at some point the repulsive force from these counterions balances the electrostatic attraction between the surfaces.

## 2.3 INTERMOLECULAR FORCES AND AGGREGATION

Supramolecular chemistry is dominated by the host of non-covalent interactions present in molecular subunits. A simple illustration can help us understand the interplay between the various interactions discussed so far, and how this interplay leads to molecular self-assembly into nanomaterials with a specific structure. Consider the set of generic molecules shown in Figure 2.14. For simplicity, only characteristics emphasizing interactions are shown such as ionic moieties, hydrophobic regions, dipoles, and hydrogen bonding groups. The molecules represent molecular building blocks and the aggregation of these individual blocks will be affected by intermolecular interactions.

The organization of the building blocks into more complex structures will largely be driven by a minimization of energy and thermodynamic constraints. The latter factor, for example, can be an entropic gain due to the hydrophobic effect. Minimization of energy will be achieved by minimizing unfavorable interactions such as bringing two like-charged moieties to the same vicinity. Possible ways the molecular building blocks may assemble are shown in Figure 2.14(b). Self-assembly will lead to a three-dimensional aggregate, but if the assembly



**Figure 2.14** (a) Representation of four molecular building blocks containing various interacting functionalities. (b) Possible aggregation patterns driven by (i) like-charge repulsion, (ii) dipole-dipole and H-bonding interactions, (iii) opposite-charge attraction, and (iv) strong hydrophobic interactions. Hydrophobic interactions probably play a role in all of these aggregates. (c) Strong substrate-molecule interactions cause the molecules to tilt in order to minimize like-charge interactions.

is occurring on a surface, then a two-dimensional aggregate will form. A planar support will also impose some restrictions on the exact structure. For example, in Figure 2.14(c), a supported monolayer is forced into a tilted and/or staggered state because of strong substrate-molecule interactions and intermolecular head group repulsions.

A thermodynamically stable structure will be formed based on the conditions of the assembly process (temperature, the presence of a surface, pH, concentration, etc.). If the structure is formed in a solvent such as water, it is important to appreciate that the molecules comprising the aggregate may be in dynamic equilibrium with “free monomers” in solution. One consequence of this is that the aggregate size and shape may



change with monomer concentration and other conditions such as pH, temperature, and salt concentration.

Up until now, we have ignored direct electron coupling between neighboring molecules in a self-assembled aggregate. Although induced dipole effects are electronic in origin, the molecules may be relatively far apart so that electrons are still localized on each molecule. When one is confronted with a dense aggregate comprised of molecules that are essentially touching one another, the electronic characteristics of the material may change. For example, molecular building blocks may result in a nanostructure in which the electrons are delocalized over the entire aggregate. This will change the material's optical and electronic properties, which will be determined by the shape and size of the aggregate. The following sections provide some background in basic electronic structure pertaining to electron delocalization and the effect of size.

## 2.4 SIMPLE MODELS DESCRIBING ELECTRONIC STRUCTURE

Electrons interact with radiation, and this interaction is responsible for the absorption and emission of radiation. Phenomena such as fluorescence, phosphorescence, and photo-electricity depend on how light interacts with molecules. This interaction can be exploited to gain information about molecular structure (the basis of spectroscopy). Spectroscopy and the nature of light-matter interactions are covered in Chapter 4. Here some pertinent elements of electronic structure are covered. It is assumed that the student has a general chemistry level grasp of Lewis structures, molecular orbital (MO) theory of simple molecules, and a quantum mechanical interpretation of light (photons) and electronic structure (energy levels). One important equation worth recalling is the Planck's equation, which relates energy between two energy levels ( $\Delta E$ ) to the wavelength ( $\lambda$ ) of light absorbed or emitted as a result of an electronic transition between these two energy levels (Equation 2.24).

$$\Delta E = h\nu = \frac{hc}{\lambda} \quad (2.24)$$

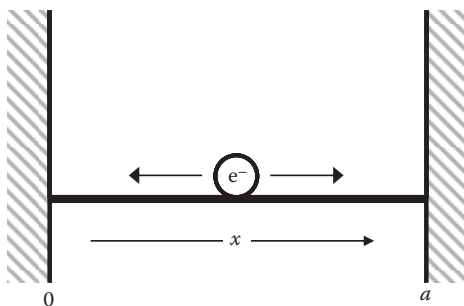
The wavelength and frequency ( $\nu$ ) of light is related by  $\lambda = c/\nu$ , where  $c$  is the speed of light in a vacuum ( $2.998 \times 10^8 \text{ ms}^{-1}$ ) and  $h$  in Equation 2.24 is the Planck's constant ( $6.626 \times 10^{-34} \text{ Js}$ ).

### 2.4.1 The Particle in a Box Model

Chemical reactivity and physical phenomena such as the absorption of light is largely determined by the electronic structure in molecules. Since electronic energy levels are quantized, Equation 2.24 provides the wavelength (or color) of light absorbed or emitted due to transitions between such levels. For example, electron configurations of atoms and molecules provide an excellent explanation of such observations. However, in many cases an electron (or electrons) is free to move within a certain region of space. This region may actually be an aggregate that has nanoscale dimensions, such as a metal nanoparticle. Metals and conjugated molecules are other examples where electrons are not restricted to the individual nuclei, but rather are delocalized over a larger region of space.

The particle in a box model is a quantum model that describes the energy levels available to an electron confined to a well-defined region of space. For simplicity we will consider this region to be a line. The line represents a one-dimensional region with infinite walls (known as potential barriers) that keep the electron confined on the line (Figure 2.15). The electron does not have the energy to jump the potential barrier, and so the probability that it is present beyond the regions 0 and  $a$  is zero. The quantum-mechanical solution of such a model tells us that the electron can only have certain values of energy, and the exact energy depends on the quantum number  $n$ , which can take values 1, 2, 3, 4, and so on. These values of energy are given by Equation 2.25, where  $a$  is the length of the line, and  $m$  is the mass of the electron.

$$E = \frac{h^2 n^2}{8ma^2} \quad (2.25)$$



**Figure 2.15** An electron confined on a line between 0 and  $a$ . At points 0 and  $a$ , the infinite potential barriers prevent the electron from crossing over.

### Example 2.3 Calculating the Energy of an Electron in a One-Dimensional Nanoscale Region

Consider an electron that is free to move along a line of length 200 nm. What is the ground state energy of the electron? What is the energy of the electron in the  $n = 3$  state? What is the effect on the spacing between neighboring energy levels if the length of the line increased from 200 nm to 300 nm?

*Solution* The ground state energy represents the lowest energy the electron can have. This is the case when  $n = 1$ . The mass of the electron is  $9.109 \times 10^{-31}$  kg and  $h = 6.626 \times 10^{-34}$  Js. Using Equation 2.25,

$$\begin{aligned} E &= \frac{h^2 n^2}{8ma^2} = \frac{(6.626 \times 10^{-34} \text{ Js})^2 (1)^2}{8(9.109 \times 10^{-31} \text{ kg})(200 \times 10^{-9} \text{ m})^2} \\ &= \frac{4.390 \times 10^{-67} \text{ J}^2 \text{ s}^2}{2.915 \times 10^{-43} \text{ kg m}^2} = 1.506 \times 10^{-24} \text{ J} \end{aligned}$$

(note: 1 J = 1 kgm<sup>2</sup>s<sup>-2</sup>)

When  $n = 3$ , the energy is

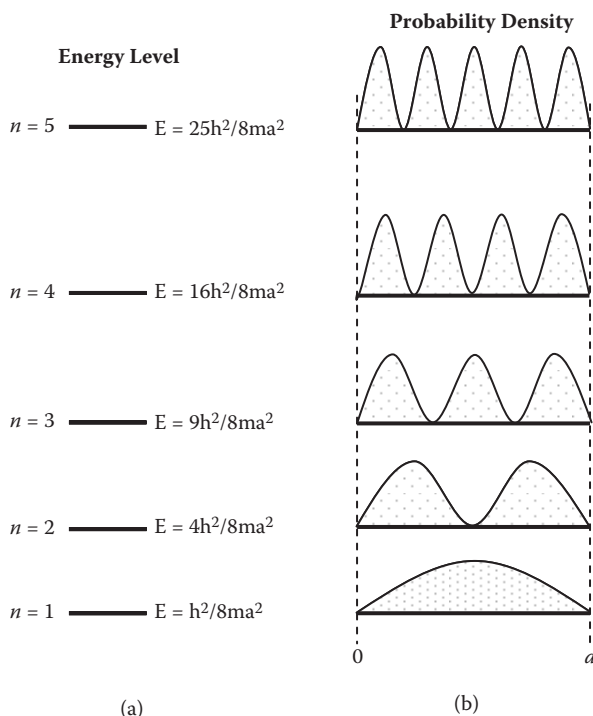
$$\begin{aligned} E &= \frac{h^2 n^2}{8ma^2} = \frac{(6.626 \times 10^{-34} \text{ Js})^2 (3)^2}{8(9.109 \times 10^{-31} \text{ kg})(200 \times 10^{-9} \text{ m})^2} \\ &= \frac{3.951 \times 10^{-66} \text{ J}^2 \text{ s}^2}{2.915 \times 10^{-43} \text{ kg m}^2} = 1.355 \times 10^{-23} \text{ J} \end{aligned}$$

The difference between neighboring energy levels ( $\Delta E$ ) is given by

$$\Delta E = E_2 - E_1 = \frac{h^2}{8ma^2} (n_2^2 - n_1^2)$$

The subscripts 1 and 2 denote lower and upper energy levels, respectively. If  $a$  increases from 200 nm to 300 nm,  $\Delta E$  will decrease by  $2^2/3^2$  or about 45%.

In the particle in a box model the electron does not exist as a discrete particle moving along the line. Rather it resembles a standing wave whose exact form depends on the value of  $n$ . The standing waves can be viewed as a smear of electron density, with regions of high and low electron probability. Regions of zero electron probability are known as nodes. Figure 2.16 shows an energy level diagram for the first few energy levels based on Equation 2.25 and representations of the

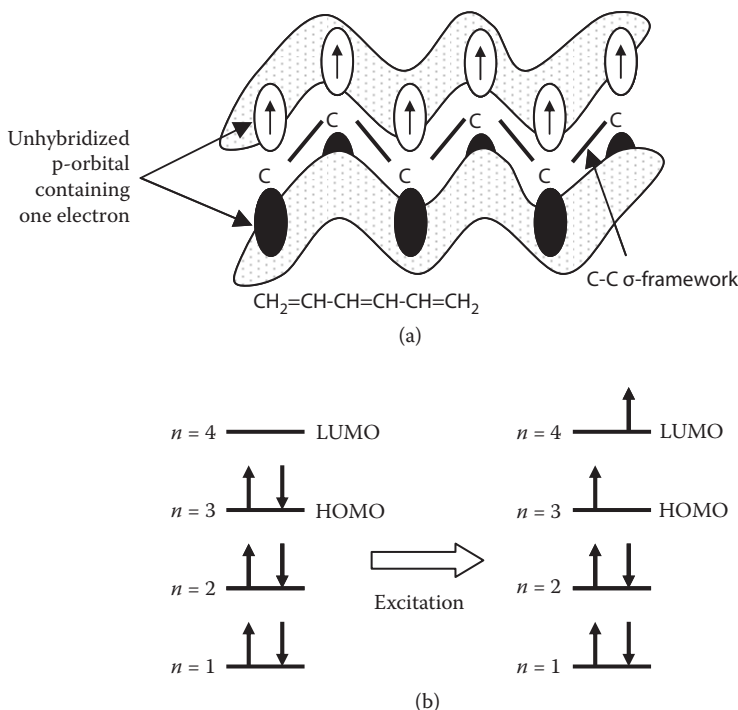


**Figure 2.16** (a) The allowed energy levels of an electron confined on a line between  $0$  and  $a$ . The energy values are given as a function of the quantum number  $n$ . As  $n$  increases, the spacing between the energy levels increases. (b) The corresponding probability density as a function of the quantum number  $n$ . The regions of zero probability represent nodes. The number of nodes is equal to  $n - 1$  (ignoring the zero probability at  $0$  and  $a$ ).

corresponding electron density for each level. If the potential barriers were not infinite, the quantum mechanical solution tells us that the electron can actually reside beyond the line region, a phenomenon known as quantum mechanical tunneling. However, the probability density in the regions beyond the line decays very rapidly.

The particle in a box model is arguably the simplest quantum mechanical model describing the energy and probability density of an electron. Despite its simplicity it has been extremely valuable in describing the absorption properties of simple molecules in which the electron is freely moving along a line. Such molecules can be linear carbon chains containing alternating single and double bonds (conjugation) in which electrons are delocalized along the entire chain.

It must be pointed out that linear conjugated molecules have more than one free electron moving along the chain.



**Figure 2.17** (a) Overlap of unhybridized p-orbitals on each carbon atom produces a  $\pi$  molecular orbital resulting in the delocalization of electron density along the hexatriene chain. (b) The particle in a box model as applied to the hexatriene molecule results in an energy-level diagram showing electron pairs in three levels. Excitation of an electron to the  $n = 4$  occurs by the absorption of energy.

For example, hexatriene has 6 free electrons as shown in Figure 2.17. Each carbon atom is  $\text{sp}^2$  hybridized, and overlap of these orbitals forms the  $\sigma$  backbone of the carbon chain, including all the C-H bonds. The electrons due to  $\sigma$ -bonding are localized between the two atoms forming the bond. The particle in a box model does not apply to these electrons. However, each carbon atom has an unhybridized p-orbital containing one electron. The p-orbitals on all the carbon atoms are able to overlap forming the  $\pi$ -framework [shown by the shaded regions in Figure 2.17(a)]. The 6  $\pi$ -electrons are delocalized along the chain and can be described using the particle in a box model.

In order to deal with the 6  $\pi$ -electrons in hexatriene, an energy-level diagram is constructed like that shown in Figure 2.17(b). The energy levels are filled up with the appropriate number

of electrons such that each level contains a maximum of two electrons. The Pauli Exclusion Principle introduced in general chemistry tells us that only two electrons can enter a given energy level, and this pairing occurs with the electrons having opposite spins to each other. In Figure 2.17, the 6 free electrons in hexatriene enter the first three energy levels. Higher energy levels ( $n > 3$ ) are empty when the molecule is in the ground state. The  $n = 3$  level is known as the HOMO (highest occupied molecular orbital) level and the  $n = 4$  level is known as the LUMO (lowest unoccupied molecular orbital) level. An electron in the  $n = 3$  state can absorb a photon of energy and as a result enter the  $n = 4$  level. This is known as the HOMO-LUMO electronic transition. Of course, the photon causing the excitation must have an energy that is equal to the HOMO-LUMO energy gap.

The particle in a box model can be expanded to describe the electron moving in both a two-dimensional rectangular region (Equation 2.26) and a three-dimensional cubical region (Equation 2.27).

$$E_{2D} = \frac{h^2}{8m} \left( \frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} \right) \quad (2.26)$$

$$E_{3D} = \frac{h^2}{8m} \left( \frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2} \right) \quad (2.27)$$

In these equations  $a$ ,  $b$ , and  $c$  represent the length, width, and height of a cube (or just  $a$  and  $b$  for a rectangle), and the  $x$ ,  $y$ , and  $z$  subscripts denote quantum numbers in the three different directions of a Cartesian coordinate system. These three quantum numbers independently assume values of 1, 2, 3, 4, and so on.

An interesting solution arises when dealing with an electron moving around a ring. The energy levels are given by Equation 2.28.

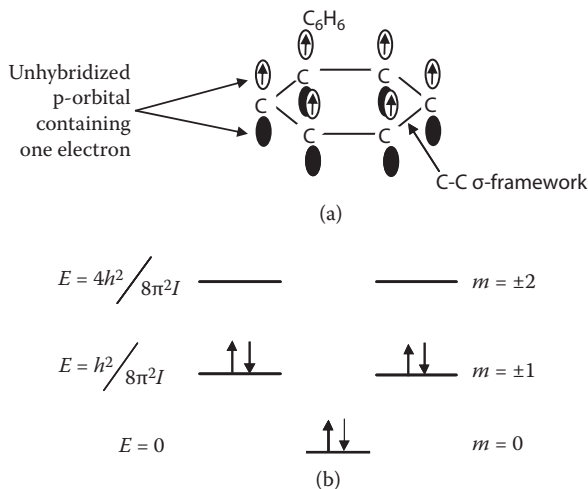
$$E = \frac{h^2}{8\pi^2 I} m^2 \quad (2.28)$$

$$I = m_e r^2 \quad (2.29)$$

$I$  is the moment of inertia of the electron going around the ring (Equation 2.29),  $r$  is the radius of the ring, and  $m_e$  is the mass of the electron. The quantum number  $m$  can take on values of 0,  $\pm 1$ ,  $\pm 2$ ,  $\pm 3$ , and so on. This means that when  $m$  is 1, there are two energy levels with the same energy. We say that the electron in the  $m = 1$  state is twofold degenerate. In general,  $+m$  and  $-m$  represent a doubly degenerate energy state.

This model can be used to describe the electronic structure of benzene. There are six  $\pi$  electrons in benzene, one in each unhybridized p-orbital on each carbon atom [Figure 2.18(a)]. These six electrons are delocalized around the ring and are regarded as free electrons. Figure 2.18(b) shows an energy level diagram based on Equation 2.28. We can place two of them into the  $m = 0$  level and four into the  $m = \pm 1$  level. The first electronic transition would be the  $1 \rightarrow 2$  transition, and the energy associated with this transition is given by Equation 2.30.

$$\Delta E = \frac{h^2}{8\pi^2 I} (2^2 - 1^2) \quad (2.30)$$



**Figure 2.18** (a) Overlap of unhybridized p-orbitals on each carbon atom produces a  $\pi$  molecular orbital resulting in the delocalization of electron density along the benzene ring. (b) The particle in a box model as applied to the benzene molecule results in an energy level diagram showing electron pairs in three levels. The two energy levels corresponding to  $m = \pm 1$  are doubly degenerate.

### Example 2.4 Estimating the Size of the Benzene Ring

Benzene absorbs light of wavelength  $\sim 250$  nm. Estimate the radius of the benzene ring.

*Solution* We first change the absorption wavelength into the corresponding energy using Equation 2.24.

$$\Delta E = \frac{hc}{\lambda} = \frac{(6.626 \times 10^{-34} \text{ Js})(2.998 \times 10^8 \text{ ms}^{-1})}{(250 \times 10^{-9} \text{ m})} = 7.946 \times 10^{-19} \text{ J}$$

Assuming that the  $1 \rightarrow 2$  transition is associated with the absorption wavelength of 250 nm, we rearrange Equation 2.30 to determine the moment of inertia.

$$\begin{aligned} I &= \frac{h^2(2^2 - 1^2)}{\Delta E(8\pi^2)} = \frac{(6.626 \times 10^{-34} \text{ Js})^2(3)}{(7.946 \times 10^{-19} \text{ J})(8\pi^2)} = 2.102 \times 10^{-50} \text{ Js}^2 \\ &= 2.102 \times 10^{-50} \text{ kgm}^2 \end{aligned}$$

(note:  $1 \text{ Js}^2 = 1 \text{ kgm}^2$ )

Finally, the radius can be estimated using Equation 2.29.

$$r = \sqrt{\frac{I}{m_e}} = \sqrt{\frac{2.102 \times 10^{-50} \text{ kgm}^2}{9.109 \times 10^{-31} \text{ kg}}} = 1.519 \times 10^{-10} \text{ m} = 0.152 \text{ nm}$$

This is close to the experimentally measured value of around 0.25 nm.

The particle in a box model can be used to describe the energy of an electron confined within a three-dimensional region of nano-scale dimension. This is known as quantum confinement, and examples involving quantum confinement in solid nanoparticles are given in Chapter 5.

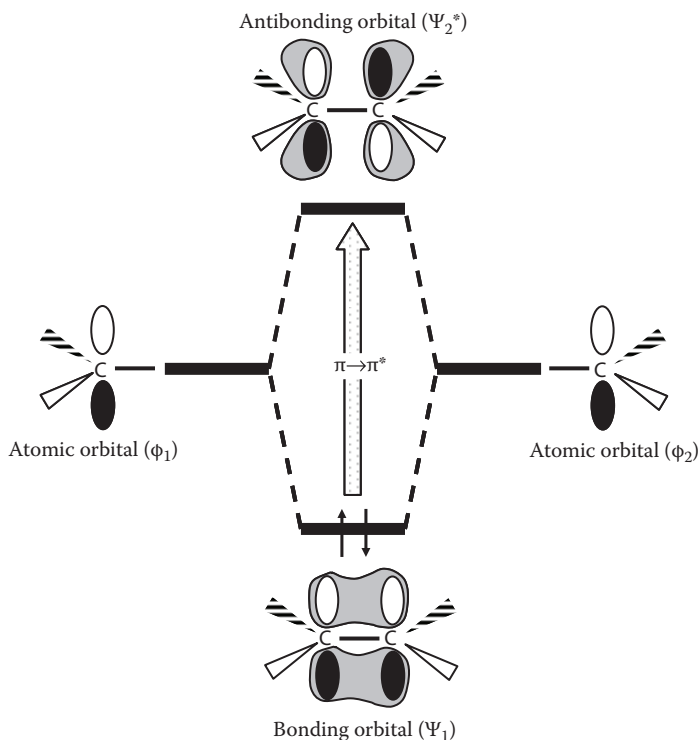
### 2.4.2 Conjugation in Organic Molecules

The free electron model can be applied to conjugated organic molecules. In fact, the model can help us explain why the absorption wavelength decreases as the length of the molecule increases. We will use the term *conjugation length* to describe the length of the alternating double- and single-bonded hydrocarbon chain in which the electrons are delocalized over. First let's review some background on the structure of conjugated organic molecules.

As briefly mentioned in the last section, a conjugated system has three  $sp^2$  hybridized orbitals on every carbon atom,



which form covalent bonds with nearby atoms. This accounts for the  $\sigma$ -bonding in the molecule. The leftover unhybridized  $p_z$  orbital combines with other  $p_z$  orbitals to form a delocalized  $\pi$  MO that spans the length of the molecule. For example, ethylene has two atomic  $p_z$  orbitals,  $\phi_1$  and  $\phi_2$ . From these two atomic orbitals, two MOs,  $\Psi_1$  and  $\Psi_2^*$ , are formed by taking linear combinations. The bonding MO,  $\Psi_1$ , results from the in-phase combination of the wave functions of the two  $p$  orbitals, whereas the antibonding orbital,  $\Psi_2^*$ , results from the out-of-phase combination. The overlap results in two new MOs—one bonding orbital with an energy lower than either of the original  $p$  orbitals and one antibonding orbital with an elevated energy. The relative energies of these MOs are illustrated in Figure 2.19. It must be pointed out that only the  $\pi$  MOs are shown in the figure. These MOs result from the overlap

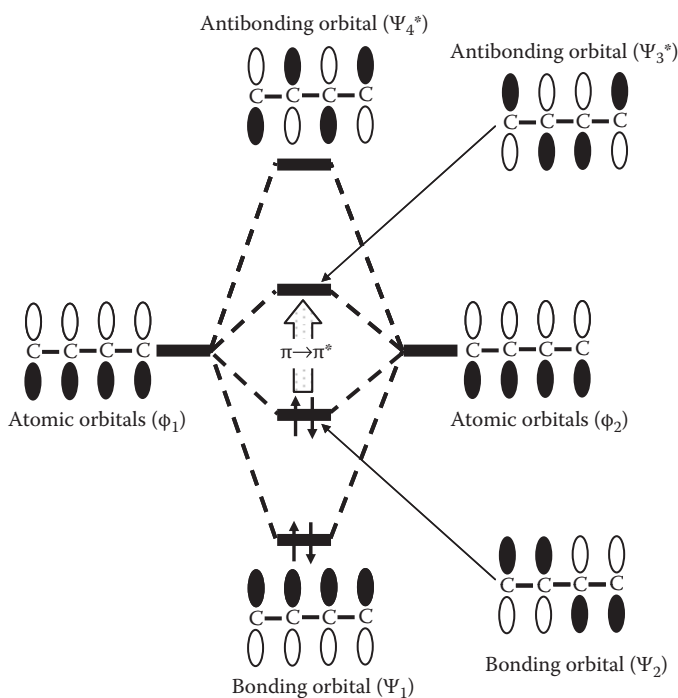


**Figure 2.19** A partial MO energy-level diagram for ethylene emphasizing the  $\pi$  bonding ( $\Psi_1$ ) and antibonding ( $\Psi_2$ ) MOs. MOs are constructed by the linear combination of the AOs of ethylene, in this illustration the unhybridized  $p$ -orbitals ( $\phi_1$  and  $\phi_2$ ) on each carbon atom. The two electrons from each  $p$ -orbital are placed in the bonding MO ( $\Psi_1$ ). The MOs due to  $\sigma$ -bonding are not shown.

between unhybridized p-orbitals on the carbons. The MOs resulting from  $\sigma$ -bonds are not shown, as this is not relevant to our discussion of conjugation.

In terms of shape, the bonding MO has electron density above and below the line connecting the two carbon atoms. There is no node between these atoms. In contrast, the antibonding MO has a node between the two carbons. The antibonding MO has electron density concentrated near each of the two carbon atoms, but zero electron density between these atoms. The presence of a node means no bond, thus the term antibonding MO.

The simple two p-orbital overlap scheme can be expanded to conjugated systems. We start by considering 1,3-butadiene, the simplest conjugated system. Because 1,3-butadiene has four atomic orbitals, four molecular orbitals must result. Figure 2.20 shows the atomic and molecular orbitals of 1,3-butadiene and

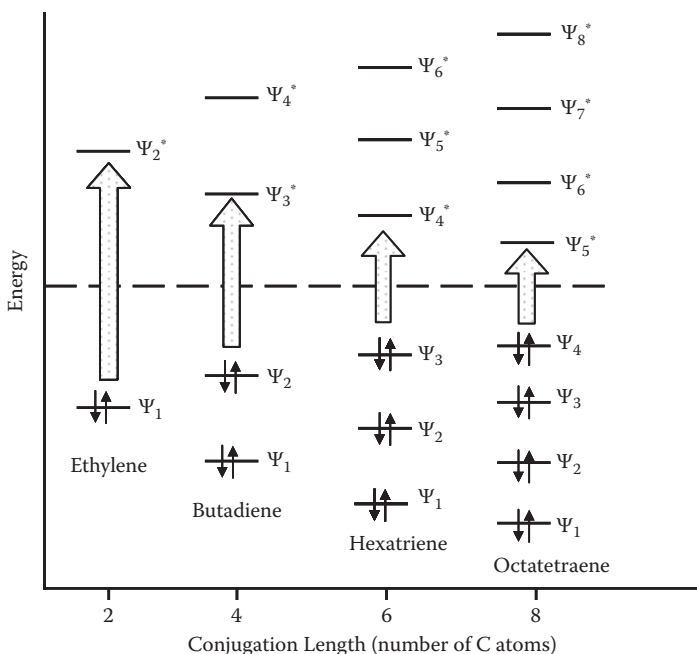


**Figure 2.20** A partial MO energy-level diagram for butadiene emphasizing the  $\pi$  bonding ( $\Psi_1$  and  $\Psi_2$ ) and antibonding ( $\Psi_3^*$  and  $\Psi_4^*$ ) MOs. MOs are constructed by the linear combination of the AOs of butadiene, in this illustration the unhybridized p-orbitals on each carbon atom grouped together and described as  $\phi_1$  and  $\phi_2$ . The four electrons from each p-orbital are placed in the bonding MOs ( $\Psi_1$  and  $\Psi_2$ ). The MOs due to  $\sigma$ -bonding are not shown.

the corresponding energy levels. Like ethylene, only the  $\pi$  MO are shown, which result from the overlap between unhybridized p-orbitals on the carbons. The MOs resulting from  $\sigma$ -bonds are not shown.

The important aspect of Figure 2.20 is the  $\pi \rightarrow \pi^*$  transitions. Notice that the  $\Psi_2 \rightarrow \Psi_3^*$  transition in 1,3-butadiene is much smaller than the  $\Psi_1 \rightarrow \Psi_2^*$  transition in ethylene. As the number of p orbitals increases in a conjugated system, the energy gap,  $\Delta E$ , between the HOMO and the LUMO becomes progressively smaller. Figure 2.21 shows the MO energy levels of ethylene and the first three conjugated systems: 1,3-butadiene; 1,3,5-hexatriene; and 1,3,5,7-octatetraene. The vertical arrows emphasize the HOMO-LUMO transitions. Further, the energy difference between the HOMO and LUMO determines the position of the absorption maximum ( $\lambda_{\max}$ ). Since wavelength is inversely proportional to  $\Delta E$ ,  $\lambda_{\max}$  increases with the length of the chain of the conjugated carbon atoms.

In the extreme case of a conjugated polymer, the  $\pi$  bonding orbitals and the  $\pi^*$  antibonding orbitals form broad bands separated by a small  $\Delta E$ . A long conjugated polymer is able



**Figure 2.21** The MO energy-level diagram for a series of linear conjugated hydrocarbons. The figure shows how  $\Delta E$  for the  $\pi \rightarrow \pi^*$  transition changes as the conjugation length increases.

to conduct electricity because electrons in the valence band ( $\pi$ ) can be excited into the conduction band ( $\pi^*$ ). There, they are free to move and carry charge. Because electrons in conjugated polymers must be excited in order to conduct electricity, conjugated polymers make excellent semiconductors.

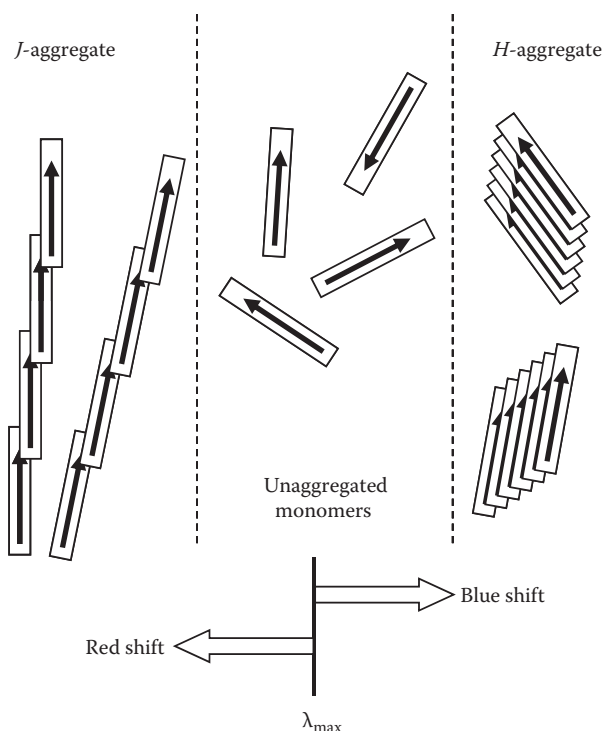
Alan Heeger, Alan MacDiarmid, and Hideki Shirakawa shared the Noble Prize in Chemistry in 2000 for their discovery and development of conductive (conjugated) polymers in the late 1960s and early 1970s. They did much of their work on polyacetylene, the simplest conjugated polymer, with alternating single and double bonds in a linear carbon chain. Polyacetylene, however, is very susceptible to photo-oxidation, and more recent studies have focused on other, more stable conjugated polymers. Derivatives of polythiophene, polyaniline, polyfluorene, and poly(phenylene vinylene) or PPV have similar conductivities to polyacetylene but are much more stable against oxidative degradation. Polymers and other macromolecules will be discussed in Chapter 5.

### 2.4.3 Aggregation and Electronic Structure

Molecular aggregates have electronic and spectroscopic properties that may be considerably different from the monomer. Consider a conjugated molecule that contains a dipole moment. We can represent this molecular building block as a rod shown in Figure 2.22, where the arrow indicates the direction of the dipole. Let's form a molecular aggregate in which individual monomers are arranged in a regular fashion. We can distinguish between two types of aggregation patterns called the *H*-type and *J*-type aggregation.

The *J*-type aggregate is a one-dimensional molecular assembly in which the dipole moments of the individual monomers are aligned parallel to the line joining their centers. This is sometimes referred to as the "end-to-end arrangement." In contrast, the *H*-type arrangement, while still a one-dimensional array, is one in which the dipole moments are aligned parallel to each other but perpendicular to the line joining their centers. This is sometimes referred to as the "face-to-face arrangement." Figure 2.22 illustrates the difference between a *J*-aggregate and an *H*-aggregate.

One of the most characteristic properties of *J*-type aggregation is that such materials absorb higher wavelength light with respect to the monomer absorption. We say that it is red-shifted in the absorption spectrum with respect to the monomer absorption (see Chapter 4). The absorption wavelength of the *H*-aggregate is a little lower (or blue-shifted) with



**Figure 2.22** Non-covalent aggregation patterns in dipolar building blocks. The *J*-aggregate is a head-to-tail arrangement and the *H*-aggregate is a head-to-head arrangement between neighboring molecules.

respect to the monomer absorption wavelength. The energy shift of the absorption wavelengths of the aggregates has been explained by exciton theory. This theory will not be covered in this text, but we can apply our understanding of conjugation length and the particle in a box models to help understand this observation.

First assume that the electron motion is along the dipole only. A *J*-aggregate can be considered a line along which free electrons can move. Since the dipoles are arranged end-to-end, the electron motion can be considered as delocalized along the entire length of the aggregate. According to Equation 2.25, the length  $a$  is much greater than the length of an individual monomer. Therefore,  $\Delta E$  is small and is inversely proportional to wavelength (Equation 2.24), and the aggregate absorbs light energy at a higher wavelength compared to a monomer. Conversely, an *H*-aggregate will have a short conjugation length and a larger  $\Delta E$  value

corresponding to the absorption of light of a smaller wavelength. The unaggregated monomer can be considered intermediate between these two extremes, and so the  $\Delta E$  value corresponding to the excitation lies between the values for the *H*- and *J*-aggregate.

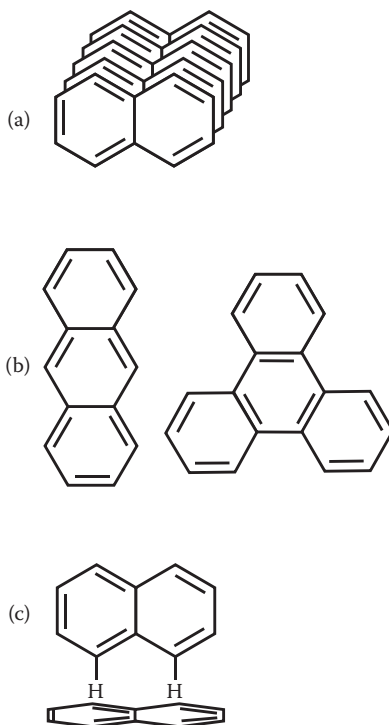
#### 2.4.4 $\pi$ - $\pi$ Stacking Interactions

We end this chapter with a brief discussion of a very weak electron coupling interaction leading to aggregation known as the  $\pi$ - $\pi$  stacking interaction. These interactions occur due to the presence of p-orbitals in conjugated ring systems such as benzene. The net effect of such an interaction is face stacking of planar rings as shown in Figure 2.23(a) for naphthalene. Although the effect is actually insignificant in small systems such as benzene, the interaction becomes stronger as the number of  $\pi$ -electrons increases. It should be pointed out, however, that even in these systems, electrostatic forces usually overcome  $\pi$ - $\pi$  stacking interactions. Nonetheless, the interaction is particularly strong in planar polycyclic aromatic molecules containing many delocalized rings. Examples include anthracene and triphenylene [Figure 2.23(b)]. For large ring systems these interactions can be so significant that they may dominate the supramolecular chemistry and determine the overall structure of the aggregate. For example, the interaction determines the growth of organic crystals composed of such polycyclic molecules.

$\pi$ -stacking interactions affect the properties of polymers, peptides, liquid crystals, and proteins. In biology,  $\pi$ -stacking occurs between adjacent nucleotides, and this adds to the stability of double-stranded DNA. A related phenomenon, called the edge-face interaction, is often observed in proteins where the hydrogen atom of one aromatic system points perpendicularly to the center of the aromatic plane of the other aromatic system [Figure 2.23(c)]. This type of interaction is thought of as related to the partial positive charge on the hydrogen atom connected to the aromatic ring.

## REFERENCES AND RECOMMENDED READING

- Israelachvili, J. *Intermolecular and Surface Forces*, 3rd ed., 2011, Academic Press, San Diego, CA, pp. 31–139. This is a classic textbook on the subject, but is probably for the more advanced student. Chapters 3 through 8 are particularly relevant to this text.



**Figure 2.23** (a)  $\pi$ - $\pi$  stacking between planar aromatic rings. (b) Structures of anthracene and triphenylene. (c) The edge-face interaction between two planar aromatic ring systems.

- Lowe, J. P. *Quantum Chemistry*, 2nd ed., 1993, Academic Press, San Diego, CA. This book is for the advanced student. Chapter 2 contains some excellent applications of the particle in a box model.
- Pavia, D., Lampman, G., and Kriz, G. *Introduction to Spectroscopy*, 2nd ed., 1996, Saunders College Publishing, Orlando. The chapter on electronic spectroscopy provides some useful background on MO diagrams of organic molecules. This book is an excellent read for those interested in the kinds of electronic excitations discussed in Section 2.4.2.
- For those interested in H- and J-aggregation, the following is excellent: Kuhn, H. and Kuhn, C. *Chromophore Coupling Effects*. In *J-Aggregation*, Kobayashi, T., Ed., 1996, World Scientific: Singapore, 1–140.

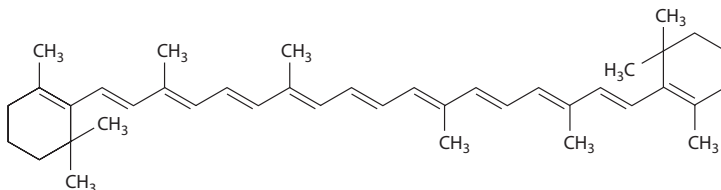
## END OF CHAPTER QUESTIONS

1.  $k_b T$  is often used to measure the strength of a given interaction, where  $k_b$  is Boltzmann's constant ( $1.38065 \times 10^{-23} \text{ J K}^{-1}$ ). Coulombic interactions are generally among the strongest and the farthest reaching of the interactions that we have discussed. Using the information given in Example 2.1, determine at what distance  $\text{Na}^+$  and  $\text{Cl}^-$  would have to be separated in order for  $V(r)_{\text{Na-Cl}}$  to be equal in magnitude to  $k_b T$  at room temperature (298K).
2. (a) Imagine that a  $\text{Ca}^{2+}$  ion interacts with an  $\text{H}_2\text{O}$  molecule in the absence of any other molecules at a distance of 0.4 nm. Plot the interaction potential energy as a function of  $\theta$ . Assume that the only interaction between the two species is an ion-dipole interaction and that the distance remains fixed. (b) At what angles does the magnitude of the potential energy reach a maximum? Why does this make sense? What do the positive and negative values of  $V$  represent?
3. Obtain an expression for the force between an ion and its induced dipole in an ion-induced dipole interaction. Is the force attractive or repulsive?
4. (a) Assume that two dipolar molecules (each with dipole moment  $\mu$ ) are in the same plane. Show that for all distances  $r$ , the interaction potential energy between the dipoles is smaller (more negative) if the two dipoles are oriented in a line rather than anti-parallel to each other. (b) The answer from 4(a) is slightly misleading because it might lead us to believe that the dipoles always prefer to orient themselves in a line rather than anti-parallel. If the dipolar molecules are anisotropic in shape (oblong in the direction of the dipole), explain why the molecules might prefer to orient themselves in an anti-parallel orientation rather than in a line.
5. Explain in your own terms why, in the absence of other interactions, a dipole-induced dipole interaction will always be attractive. Use diagrams if appropriate.
6. At what value of  $r$  (in terms of  $\sigma$ ) does the Lennard-Jones potential reach its minimum value? Mathematically, what does  $\sigma$  represent?
7. Use the particle in a box model to predict the values of  $\Delta E$  and the corresponding values of  $\lambda_{\text{max}}$  for the



conjugated linear hydrocarbons shown in Figure 2.21. Assume the C-C bond length is 154 pm and the C=C bond length is 134 pm.

8.  $\beta$ -Carotene is an organic molecule responsible for the red-orange color in plants and fruits. The molecular structure is shown below. Assuming that the length of the molecule is 2.94 nm, determine the value of  $\lambda_{\text{max}}$ . Does the result agree with the red-orange color observed for this molecule?

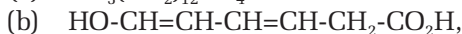


9. This question concerns the “quantum” confinement of an electron within a nano-cube. When sodium dissolves in liquid ammonia, some dissociation occurs according to the following equation:



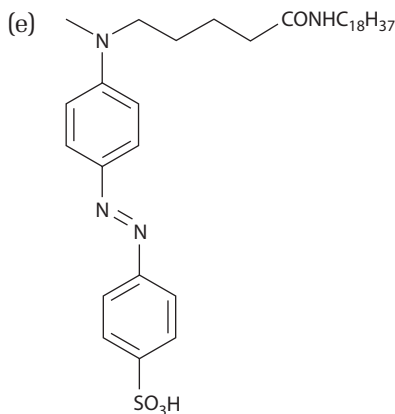
The solvated electron can be treated as a particle in a three-dimensional box. Assume that the box is cubic with an edge length of  $1.55 \times 10^{-7}$  cm, and suppose that excitation occurs in all directions simultaneously from the lowest state ( $n = 1$ ) to the first excited state ( $n = 2$ ). What wavelength, in nm, of radiation would the electron absorb? Would you expect the solution to be colored?

10. Discuss the possible types of aggregation patterns when the following water-soluble molecules self-assemble in the aqueous phase. Your discussion should include intermolecular contributions from van der Waals interactions, electronic coupling effects, and predictions about the absorption properties of the resulting aggregates. How would molecules aggregate if they were forced at the interface of a hydrophobic surface and a water phase?

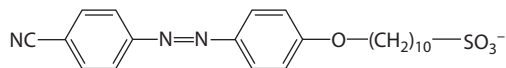
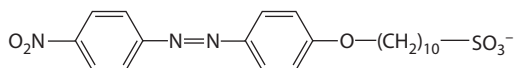
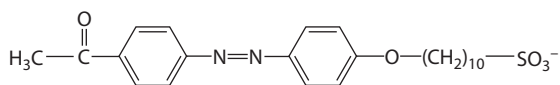




where  $X = \text{OH}$  or  $\text{NO}_2$



11. The following molecules are identical except for the structure of dipolar head groups. Place the molecules in order of increasing head group dipole moment and discuss the differences in the way these molecules aggregate into a monolayer at the air-water interface, if any difference exists.





# Three

---

## Rudiments of Surface Nanoscience

---

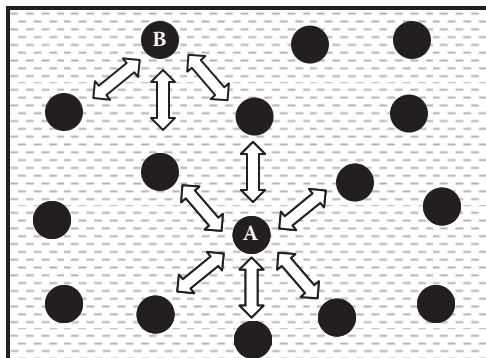
### CHAPTER OVERVIEW

Surfaces and interfaces occur everywhere in nature, from biological cells to the vast expanse of oceans. They play a key role in nanoscience since they are often used as platforms for the growth of nanomaterials. Understanding nanomaterials would therefore not be complete without covering important elements of surface science. Interfaces represent the two-dimensional plane between two different bulk phases of matter, such as oil and water. We will describe a surface as an interface where one of the bulk phases is a gas, usually air. Physical and chemical processes occurring at such regions tend to be very different from corresponding processes in the bulk phase as such processes are confined to a region in which one dimension is on the nanoscale order. For instance, a special interaction known as the hydrophobic effect influences chemistry at surfaces, often responsible for the formation of films of nanoscale thickness. In this chapter, a consideration of hydrophobicity and the surface energy of solids and liquids will lead to a discussion of contact angles and wetting phenomena. This naturally leads to a discussion of self-assembled monolayers and adsorption phenomena. An understanding of how intermolecular interactions influence the adsorption and aggregation of molecules into nanostructures (such as micelles) is provided by considering the amphiphilic nature of surfactant molecules.

### 3.1 FUNDAMENTALS OF SURFACE SCIENCE

#### 3.1.1 The Surface Energy of Solids and Liquids

Molecules of the same type tend to experience a net attractive interaction. In liquids, this cohesive force keeps molecules



**Figure 3.1** Bulk and surface interactions between molecules in a pure phase (e.g.,  $\text{H}_2\text{O}$ ). Bulk molecule A is surrounded symmetrically by its neighbors. Surface-bound molecule B is surrounded asymmetrically by its neighbors.

close to each other, and each molecule is symmetrically surrounded by others, resulting in a zero net interaction. The picture is considerably different on the surface since molecules here are not surrounded symmetrically by others (Figure 3.1). At the surface a molecule experiences cohesive forces from others in the bulk beneath it, but there are no interactions above it. This asymmetry in force results in a net inward pull on the surface-bound molecule normal to the surface. This is the molecular basis of surface tension, which is defined as the force acting parallel to the surface and at right angles to a line of unit length on the surface. The experimental measurements of surface tension are described in Chapter 4.

The existence of surface tension is the reason why many liquids, such as water, tend to spontaneously contract and minimize the surface-area-to-volume ratio. To this end, such liquids adopt a spherical geometry in the absence of all external forces such as gravity. One can determine the work done in expanding the surface of a liquid against the surface tension forces. This work actually represents the surface free energy of the expansion of the interface.

### 3.1.2 Surface Free Energy of Adsorbed Monolayers

The measurement of surface tension is perhaps the oldest method used to characterize a nanofilm floating on the surface of water. Later in this chapter we will encounter a class of molecules known as amphiphiles, which are capable of self-assembling into nanostructures of remarkable complexity.

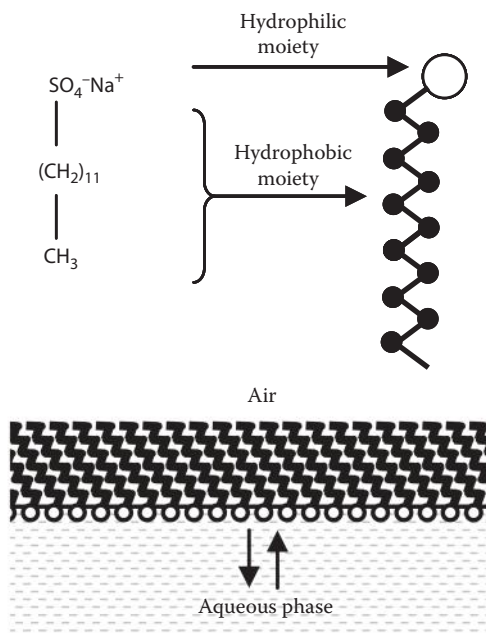
Amphiphilic molecules, such as lipids and surfactants, have a tendency to accumulate on the surface of water. The driving force behind this self-assembly process is a combination of factors including the lack of solubility of the hydrocarbon portion of the molecule in water and the tendency of the charged or polar head group to point toward the aqueous phase. The net result is the formation of a film composed of a monolayer of highly oriented molecules. The closely packed monolayer is in dynamic equilibrium with individual molecules in the bulk aqueous phase, and so the packing density of the film increases as the concentration of molecules in the bulk phase increases. However, structural and thermodynamic constraints often limit the packing density, and a saturated monolayer is usually formed when the bulk concentration is sufficiently high. The thickness of such a film is approximately the length of the amphiphilic molecule (nanoscale). Due to this short dimension, the monomolecular film is often described as a two-dimensional nanoassembly.

Figure 3.2 shows a monolayer of the surfactant sodium dodecyl sulfate (SDS) confined to the air-water interface. Later in the book, we show that nanofilms at the air-water interface are of vital importance to biological processes, and their properties can be exploited in a number of important applications. Thus, it is often crucial to know how well the film is packed within both the monolayer and the cross-sectional area occupied by each molecule on the liquid surface. These factors often determine the physicochemical properties of the film.

Surface tension measurements provide a simple but powerful method for determining a variety of characteristics of adsorbed monolayers, including the density of the monolayer. Surface tension measures the stability of a surface. If the surface is relatively unstable, it has a large surface tension value and is considered a “high energy surface.” For example, liquid water has a high surface energy because the molecules prefer not to be on the surface, but rather to be surrounded completely by other water molecules and interact with them through energetically favorable intermolecular hydrogen bonding.

A more rigorous way to define surface tension is to say that it is the free energy required to transport a molecule from the bulk phase to the surface and hence expand the area of the surface ( $dA$ ). This surface free energy ( $dG_{\text{surf}}$ ) is given by

$$dG_{\text{surf}} = \gamma dA \quad (3.1)$$

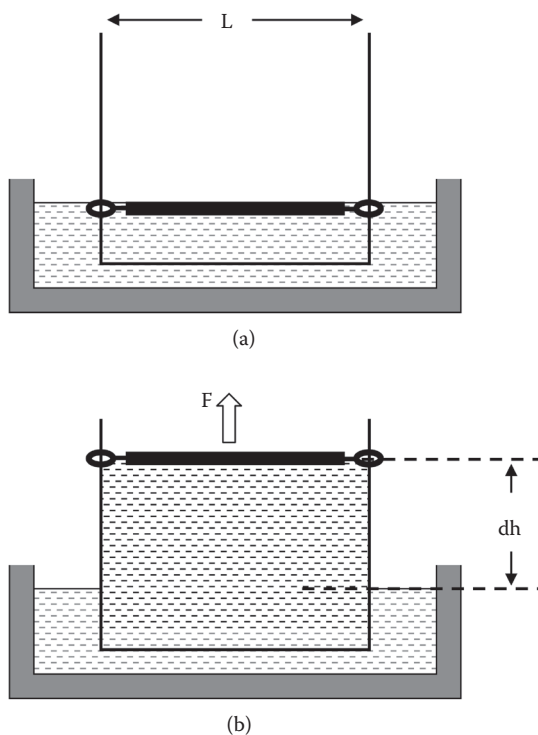


**Figure 3.2** A monolayer of the anionic surfactant sodium dodecylsulfate (SDS) at the aqueous-air interface. The monolayer is in equilibrium with SDS molecules in the bulk aqueous phase. The hydrophobic moieties of SDS are pointing away from the aqueous phase and the polar headgroups are buried in the aqueous phase. This orientation of the SDS molecules on the surface stabilizes the air–water interface.

where the proportionality constant  $\gamma$  is the surface tension. The units of surface tension are energy per unit area,  $\text{J m}^{-2}$  (or since energy can be thought of as applied force multiplied by the distance moved, the units are also Newton's per meter,  $\text{N m}^{-1}$ ). These dimensions are also the same as the spring constant (from Hooke's law), which essentially measures the resistance of a spring to stretch due to an applied force. Thus, we can interpret surface tension as a measure of the resistance of a surface to increase its area.

### Example 3.1 Determining the Work Done in Expanding a Liquid Film

Consider a simple experiment in which a film is withdrawn from a soap solution as shown in Figure 3.3. By noting that work done is force times displacement, and the fact that this work is the surface



**Figure 3.3** A film being withdrawn from a soap solution to a height  $dh$  by a force  $F$ , using a wire frame of length  $L$ .

free energy given by Equation 3.1, show that surface tension is  $F/2L$ . Prove that the units of surface tension are  $\text{N m}^{-1}$ .

*Solution* Work must be done to pull up the frame and create a film [Figure 3.3(b)].

Work done ( $dw = F dh$ ) is equal to the surface free energy ( $\gamma dA$ ), where  $A$  is the area of the surface. The soap film has two sides or surfaces, so in this example area  $A$  is actually  $2A$ . Also,  $A$  is the length  $L$  times the distance moved  $dh$ .

Therefore,  $F dh = \gamma dA = \gamma 2L dh$ .

Rearranging gives

$$\gamma = \frac{F}{2L}.$$

Since force is measured in  $\text{N}$ , the units of  $\gamma$  are  $\text{N m}^{-1}$ .



**Table 3.1** Surface Tension Values of Some Pure Liquids

Liquid	$\gamma/\text{mN m}^{-1}$		
	25°C	50°C	75°C
Water	71.99	67.94	63.57
1-Decanol	28.51	26.68	24.85
Mercury	485.48	480.36	475.23
Ethanol	21.97	19.89	—
Bromine	40.95	36.40	—
Pyridine	36.56	33.29	30.03
Toluene	27.93	24.96	21.98
Benzene	28.22	25.00	21.77

Source: *CRC Handbook of Chemistry and Physics, Internet Version 2007*, 87th ed., Taylor and Francis, Boca Raton, 2007. With permission.

The larger the surface tension, the greater the resistance to increase the surface area, which in turn fundamentally depends on the interplay between various intermolecular interactions occurring on the surface. Table 3.1 lists the surface tension values of a few common liquids at three different temperatures.

Water has one of the highest known surface tension values (about  $72 \text{ mN m}^{-1}$  at room temperature) for the reasons mentioned above. However, when amphiphilic molecules are present in water, they tend to aggregate on the surface and lower the surface tension. For example, adding a milligram of SDS to approximately 200 mL of pure water lowers the surface tension by about  $30 \text{ mN m}^{-1}$ . This decrease in surface tension occurs because the amphiphilic molecules orient themselves at the interface such that they expose their insoluble hydrophobic tails to the air while keeping the polar head group buried in the aqueous phase, thereby freeing up water molecules that would otherwise be forced to remain unfavorably at the air-water interface (Figure 3.2). These conditions stabilize the surface and consequently lower the surface tension. The degree to which the surface tension decreases depends on the structure of the amphiphile and the packing density of the resulting monolayer. If the concentration of the amphiphiles is low, then we can expect that the number of molecules on the surface is also relatively small. In a way, this dilute surface approximates a gaseous phase in which individual molecules are far apart and are free to move over the surface in a random fashion. As the concentration increases, however, the packing of the amphiphilic

molecules at the surface becomes denser and the surface tension consequently decreases. This behavior continues with rising concentration until the point at which a saturated monolayer is formed. Beyond this concentration the surface tension value does not change. The phase behavior of amphiphiles at the air-water interface is discussed in Chapter 5.

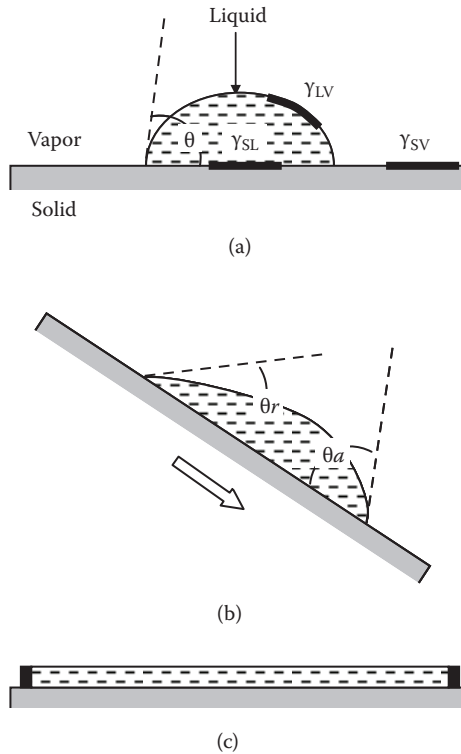
### 3.1.3 Contact Angles and Wetting Phenomena

When a drop of water is placed on a planar solid surface, it may at one extreme completely spread to cover the entire surface, or at the other extreme form a spherical droplet on the surface. These situations represent either complete wetting or complete de-wetting. Usually the degree of wetting is intermediate between these extremes and depends largely on the interfacial energy (or surface tension) between the liquid and the solid surface. The contact angle is the angle at which the liquid-vapor interface meets the solid surface (Figure 3.4). If the planar surface is horizontal and the droplet is not moving, this angle is called the static contact angle [Figure 3.4(a)]. If the liquid is in motion because the surface is tilted, then we can identify two “dynamic” contact angles, the advancing contact angle and the receding contact angle [Figure 3.4(b)]. This picture is similar to a raindrop running down the surface of a window. Usually the advancing contact angle is much larger than the receding angle, and the difference between the two values is called contact angle hysteresis.

Figure 3.4(a) shows a non-wetting drop (conventionally called a “sessile drop”) on a planar solid surface making a contact angle,  $\theta$ , which must be greater than zero. The various interfaces are described by their surface tensions: the liquid-vapor tension ( $\gamma_{LV}$ ), the solid-vapor tension ( $\gamma_{SV}$ ), and the solid-liquid tension ( $\gamma_{SL}$ ). The Young equation provides a relationship between these various surface tensions and the static contact angle:

$$\gamma_{SV} = \gamma_{SL} + \gamma_{LV} \cos\theta \quad (3.2)$$

Since  $\gamma_{LV}$  and  $\theta$  can easily be measured (see Chapter 4), the value of  $\gamma_{SV} - \gamma_{SL}$  can be determined. If the liquid completely wetted the solid ( $\theta = 0^\circ$ ), then the value of  $\gamma_{SL} = 0$ . This is a hypothetical situation in which there is no “tension” between these two phases. In fact a plot of  $\cos\theta$  versus  $\gamma_{LV}$  is linear with a negative slope. The line can be extrapolated to the value  $\cos\theta = 1$  and the corresponding surface tension measured (on the x-axis).  $\cos\theta = 1$  corresponds to a contact angle of  $0^\circ$ , or complete wetting. The corresponding surface tension is called the critical



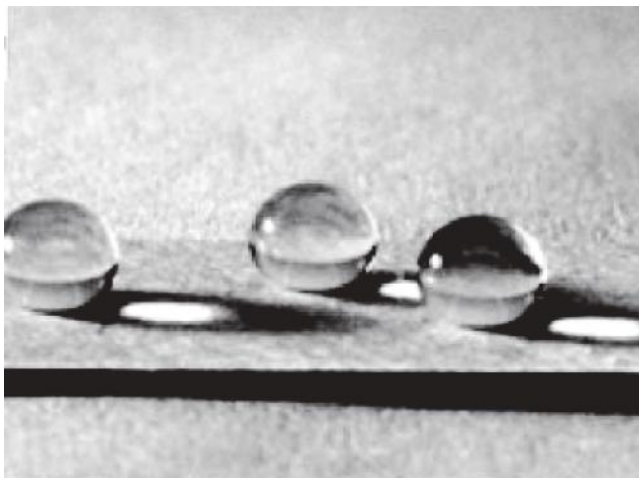
**Figure 3.4** A liquid drop on a solid surface. (a) The drop is non-wetting with a contact angle  $\theta$ . (b) The drop is running down the solid with advancing ( $\theta_a$ ) and receding ( $\theta_r$ ) contact angles. (c) Complete spreading. The solid dark lines in (a) highlight the various interfaces with different surface tensions: the liquid-vapor tension ( $\gamma_{LV}$ ), the solid-vapor tension ( $\gamma_{SV}$ ), and the solid-liquid tension ( $\gamma_{SL}$ ).

surface tension ( $\gamma_c$ ). Since under these conditions of complete wetting  $\gamma_{SL} = 0$ , the Young equation tells us that  $\gamma_c = \gamma_{SV}$ .

It is worth noting that surfaces can be hydrophobic (water hating) or hydrophilic (water loving). The surface of glass, for example, is considered hydrophilic and Teflon hydrophobic. Water will not spread on hydrophobic surfaces, but will instead form a droplet with a relatively large contact angle. Conversely, water will spread on a hydrophilic surface since it has a strong affinity for that surface, forming a thin film with a very small contact angle.

### 3.1.4 Nanomaterials and Superhydrophobic Surfaces

If the contact angle of a sessile drop approaches  $180^\circ$ , the drop essentially adopts its spherical geometry on the surface and moves around much like a frictionless bearing on the

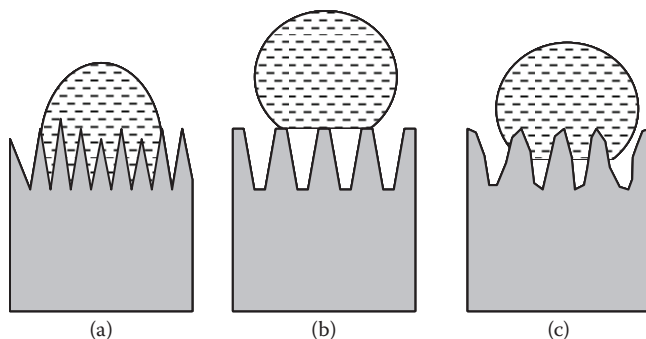


**Figure 3.5** A sessile drop on a superhydrophobic surface. (Reproduced with permission from Ma, Hill, Lowery, Fridrikh, Rutledge. “Electrospun Poly(Styrene-block-dimethylsiloxane) Block Copolymer Fibers Exhibiting Superhydrophobicity.” *ACS Langmuir* 2005, 21:5549–5554.)

surface. Figure 3.5 shows a spherical droplet of water on a hydrophobic surface. In order for a water droplet to behave this way, the surface needs to be one that minimizes the area of the solid-water interface, or a “superhydrophobic” surface. Creating such surfaces has been a fascinating challenge in nanoscience.

Superhydrophobic surfaces are not new technology by any means—scientists have been experimenting with superhydrophobicity for nearly a century, and they have been entranced by the superhydrophobicity of natural materials like the lotus leaf for far longer. However, new developments as well as new potential applications for their use have spurred the popularity of superhydrophobic surfaces. It is important, therefore, to understand the basic types of superhydrophobic surfaces, their specific characteristics, and the potentials for the use and development of each method of fabricating these surfaces. All of these concepts were outlined in the popular review article, “Progress in Superhydrophobic Surface Development,” by Paul Roach, Neil Shirtcliffe, and Michael Newton, 2008.

First, a basic understanding of the mechanism behind superhydrophobicity is important in understanding its applications. There are two basic states of superhydrophobic surfaces, each governed by separate equations and having separate characteristics. These states are illustrated in Figure 3.6. The



**Figure 3.6** The various wetting states on rough surfaces: (a) Wenzel state, (b) Cassie-Baxter state, and (c) intermediate wetting.

first type was outlined by Wenzel in 1936 and describes a wetting state in which water rests upon a surface whose morphology has been altered so that in a given area, water is in contact with more surface than if the surface were completely flat. Wenzel states are described by the equation

$$\cos\theta^W = r\cos\theta \quad (3.3)$$

where  $\theta$  is the contact angle on an unmodified surface,  $\theta^W$  is the Wenzel contact angle (on the rough surface), and  $r$  is the ratio of the actual surface area of the substrate to the projection of that surface onto a horizontal surface. In other words,  $r$  is the ratio of the actual surface area to what the surface area would be if the substrate were completely smooth. This equation essentially states that Wenzel wetting increases the contact angle of a drop of water by creating more hydrophobic surface with which the water can interact. However, if the surface is hydrophilic ( $\theta < 90^\circ$ ) then Wenzel wetting actually increases the hydrophilic properties of the surface. Also, because water is present between perturbations, water in a Wenzel state is less likely to roll off the substrate than on a flat surface of the same material. Thus, Wenzel wetting relies on two factors—an already hydrophobic substrate, and an increase in the surface area of that substrate.

The other major type of wetting was discovered by Cassie and Baxter in 1944 and describes a wetting state in which water rests upon nano- or micro-sized “pins” with air spaces in between. Unlike Wenzel wetting, Cassie-Baxter states have water not touching a significant portion of the substrate, but rather relying on the air gaps in between the substrate to form

a superhydrophobic surface. In fact, Cassie-Baxter surfaces can be fabricated with hydrophilic substrates and still display superhydrophobicity. Unlike Wenzel surfaces, water in Cassie-Baxter states rolls off surfaces much more easily than on a flat surface of the same material. These properties are reflected in the equation governing Cassie-Baxter wetting:

$$\cos\theta^C = \phi_s(\cos\theta) + (1-\phi_s)\cos\theta_x \quad (3.4)$$

where  $\phi$  represents the fraction of surface present at the top of the protrusions (where the water is in contact with the substrate),  $(1-\phi)$  represents the fraction of air gaps, and  $\theta_x$  is the contact angle over the air gaps, which is approximated at  $180^\circ$ , and  $\theta^C$  is the Cassie-Baxter contact angle (on the rough surface). Thus, there are two ways to increase superhydrophobicity of a Cassie-Baxter surface—either increase the value of  $\theta$  by increasing the inherent superhydrophobicity of the substrate, or decrease  $\phi_s$  by making bigger air gaps.

Interconversion between these two states is possible, but there is an energy barrier to overcome much like any transition state between two energy minima. This phenomenon comes into play when Cassie-Baxter surfaces display Wenzel wetting, which is possible if water falling from a great distance (i.e., rain) is forced into the crevices between roughness features. The transition from one state to another is important when discussing some methods of creating superhydrophobic surfaces.

The final important feature of superhydrophobic surfaces that must be mentioned before a more thorough description of the specifics is the advantage of multiple-scale roughness, in which perturbations range from the nanoscale to the microscale. Recent study has shown this kind of surface to increase the ease with which drops roll off the surface (in other words, decreasing the contact-angle hysteresis), the prevention of conversion from Cassie-Baxter to Wenzel states, and the tendency to convert from Wenzel to Cassie-Baxter states. This property is beneficial for a variety of reasons and will be revisited later.

Because the review article “Progress in Superhydrophobic Surface Development” specifically outlines the major methods of creating superhydrophobicity, we will refrain from repeating a list of specific methods and rather focus on unifying themes and methods of fabrication as well as the most salient features of each major method. One such feature belongs to the first method described by this article—fibrous hydrophobic

surfaces. This method can span anything from carbon nanotube-coated cotton to nanoscale polymer electrospinning. It is versatile, as fibrous materials are inherently rough, and coating with hydrophobic materials or nano-rods can increase this roughness. Very high contact angles have been reported using these methods. Also, some researchers have even reported surfaces having self-cleaning properties that are dirt and oil resistant while producing high contact angles and low hysteresis. This technology appears to be promising for the development of water- and contamination-resistant clothing that would essentially wash itself.

Another interesting feature that is unique to a specific fabrication method is the ability of certain hydrophobic semiconductors to be crystallized into a superhydrophobic surface that exhibits superhydrophobicity in the dark but superhydrophilicity in the light. One likely explanation for this phenomenon is that superhydrophilicity is generated by the excitation of electrons on the surface, which suggests that there are surface properties that can be changed by applying a voltage. Instead of excitation due to ultraviolet (UV) radiation, a voltage could be used, thus allowing for materials that can be dried instantly at the flick of a switch, or wetted completely with another flick.

Several overarching themes appear in this article as well. One of the most notable is the generation of fractal solids as superhydrophobic surfaces, which may be achieved in many ways. One common approach is crystal growth, in which a material (both organic and inorganic crystalline growths have been reported) is cooled or condensed into a fractal solid. Another common method is the diffusion-limited growth process, in which a material deposits so quickly onto a substrate that the growth of the film is limited by the flow of material over the substrate. This method, by randomness, creates small disturbances on the surface of the flat substrate, which then cause more particles to attach to these perturbations. The process continues until large bumps form, which then have small disturbances on their surfaces that create sub-bumps, and so on. The greatest benefit of these kinds of assemblies is that, being fractal, these surfaces have multiple-scale roughness. The main bumps can act as Wenzel-type surfaces, while the smaller bumps act as Cassie-Baxter surfaces. As is expected, very low hysteresis and very large contact angles have been reported on fractal solids. Another important benefit of fractal solids is their ease of fabrication—because these surfaces are usually created by chaotic movement of material, they require less specialized machinery and can be made with a variety

of materials. In fact, fractal superhydrophobic surfaces have been reported using semiconductive polymers, giving these particular surfaces the ability to switch properties as mentioned before.

One very common downside of fractal solids, however, is their tendency to be opaque. Their multi-scale roughness is very beneficial to their superhydrophobic properties, but because these features range from nanoscale to microscale, they interact with visible light. Many superhydrophobic surfaces need not be transparent. However, this kind of surface could very well be used for windshield coatings on cars, or as protective layers for photoactive devices such as LEDs or photovoltaic cells. All of these applications would be rendered useless if no light could penetrate to or from the surface.

This problem is not present in another type of superhydrophobic surface fabrication—uniform nanostructures. These structures can be generated in many ways, from carbon nanotube growth to lithographic etching. These structures are the opposite of fractal patterns in that, though they are sometimes random in arrangement, they are generally uniform in size. These materials can be easily generated using laser lithography as well as chemical etching to produce surface morphologies of a specific predetermined shape and pattern. Also, these surfaces can be easily “decorated” with additional coatings and features to produce increased hydrophobicity. Structures of this type have been reported to reach contact angles up to  $178^\circ$  using dodecanoic acid-coated, cobalt hydroxide nanopins. As well, significant ( $168^\circ$ ) though not quite as substantial Wenzel contact angles were reported using decorated carbon nanotubes. Perhaps most importantly, these kinds of structures are of one size and thus do not often absorb in the visible range. This feature makes these kinds of structures more viable for the kinds of surfaces previously discussed. The drawback of these kinds of materials is that they are much less effective at repelling water than fractal structures. They lack multi-scale roughness, and even methods that generate more random arrays are less hydrophobic than fractal surfaces.

What has emerged in recent years has not been the development of one superior form of constructing superhydrophobic surfaces. Rather, many different methods have developed, each with its own benefits and drawbacks. As more applications for superhydrophobic surfaces appear—clothing, photoactive devices, windows, and more—the drive is for a variety of superhydrophobic materials that can be customized to fit these multiple needs.



### 3.2 ADSORPTION PHENOMENA: SELF-ASSEMBLED MONOLAYERS

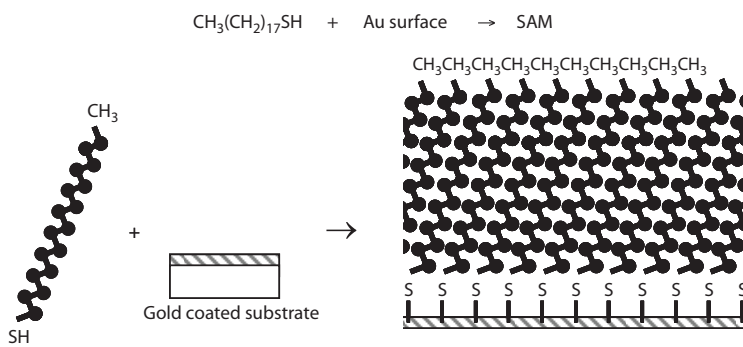
A molecule approaching a surface may experience a net attractive force and consequently become trapped or confined at the surface. Such a species is called the adsorbate, and adsorption is the physical process by which adsorbate molecules accumulate onto a solid surface. Desorption is the opposite process in which molecules leave the surface and enter the bulk phase. The solid surface in question is referred to either as the substrate or an adsorbent. The former is usually a planar surface and the latter can be a high-surface-area porous solid. The sticking of a reactant molecule to a surface was first proposed by Michael Faraday in 1834 as the initial step of a surface-catalyzed reaction. In this section we are interested in the kind of adsorption that results in a monolayer and has a thickness on the order of nanometers. The surface itself does not need to be flat but may be rough or even porous. The extent of adsorption depends not only upon the types of intermolecular forces involved (van der Waals, electrostatic, hydrogen bonding), but also on the surface area; the greater the surface area of the substrate, the greater the extent of adsorption.

The best adsorbents are those with large total surface areas, such as silica gel ( $\text{SiO}_2$ , surface area  $>1000 \text{ m}^2/\text{g}$ ) and activated carbon. Silica gel is commonly used in chromatographic columns to enhance the separation of solute mixtures by taking advantage of the different degrees of adsorption of the various components. An interesting example of an adsorption phenomenon occurs in polar stratospheric clouds in the upper atmosphere, which are highly porous and act as substrates for the adsorption of gases such as HCl.

In this and the next section we focus on the adsorption of molecules from solution or gaseous phase onto a solid substrate. However, it is important to appreciate that adsorption is a surface phenomena and can occur at any interface. For example, molecules may adsorb from an aqueous solution to the aqueous-air interface. Liquid-liquid interfaces, such as the boundary between an oil and water phase, also represent regions at which adsorption may occur. Surface adsorption plays a central role in the formation of nanomaterials. Adsorbents present themselves as platforms for the self-assembly of molecules into nanostructures. Specific examples in which nanomaterials are synthesized this way are presented in Chapter 5. At this point it is worth mentioning that solid substrates can be chemically modified so that adsorption can be selective. This

modification is known as surface functionalization and may be as simple as oxidizing a metal surface to change the surface energy (or to render it hydrophilic), so that polar molecules are spontaneously adsorbed to the surface.

Many solids have the property of adsorbing large quantities of gases and solutes from liquid solutions. This process is generally very specific both with respect to the adsorbent and the material adsorbed, driven largely by thermodynamic considerations. Adsorption is usually an exothermic process and can be divided into two kinds: chemisorption (chemical adsorption) and physisorption (physical adsorption). In general, if adsorption is specific and if large amounts of heat are liberated (greater than about  $50 \text{ kJmol}^{-1}$ ), the adsorption process is referred to as chemisorption (first proposed by Irving Langmuir in 1916). In this process, bonds are broken in the adsorbate molecules and new covalent bonds are formed between the adsorbent and the adsorbate until a complete monolayer has been established. The resulting substrate-adsorbate bond strengths range from 200 to  $500 \text{ kJmol}^{-1}$ . The chemisorbed monolayer is irreversibly bound to the solid surface and usually changes the surface properties of the solid substrate. Thus, chemisorption is an excellent method of chemically functionalizing a solid surface. As an example of chemisorption, consider the molecule octadecanethiol, or ODT, shown in Figure 3.7. The molecule contains 18 carbon atoms (17 methylene groups and one methyl group) terminated by a thiol group (SH). The thiol group is extremely reactive toward gold, resulting in a strong Au-S covalent bond. Thus, by placing a gold-coated substrate into a solution of ODT in a solvent such as chloroform,

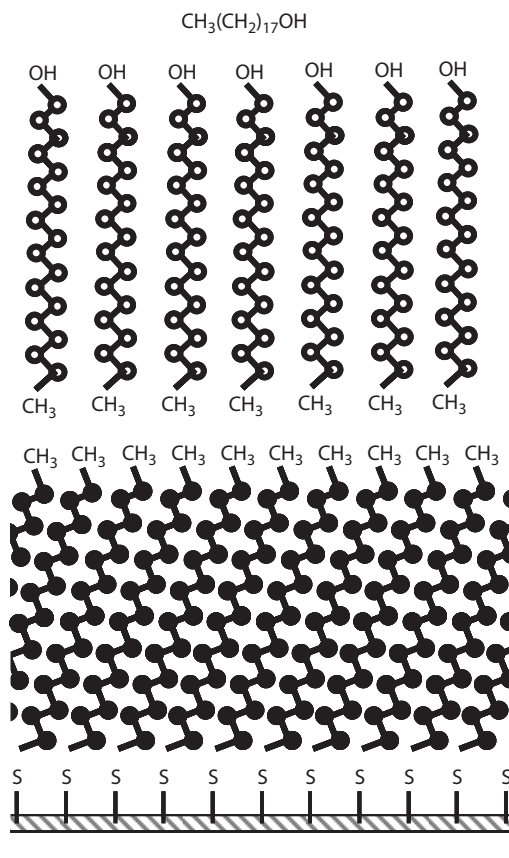


**Figure 3.7** The formation of a self-assembled monolayer by the chemisorption of octadecane thiol on a gold surface.

molecules of ODT spontaneously chemisorb to the Au surface and form a tightly packed monolayer within hours. A few features regarding this process are worth noting. First, adsorption is rapid, irreversible, and stops after a complete monolayer is formed. Second, in addition to the strong Au-S bonds that are formed, the strong van der Waals interactions between the neighboring alkyl chains in the ODT monolayer allow the molecules to pack very tightly and force the hydrocarbon backbone into an all-trans conformation (Figure 3.7). The resulting film is solid in nature, has a thickness almost the length of a fully extended ODT molecule, and is referred to as a self-assembled monolayer, or SAM. Finally, the SAM renders the solid surface hydrophobic because the surface now contains a very high density of closely packed methyl groups. Consequently, a drop of water on this surface has a large contact angle ( $>110^\circ$ ).

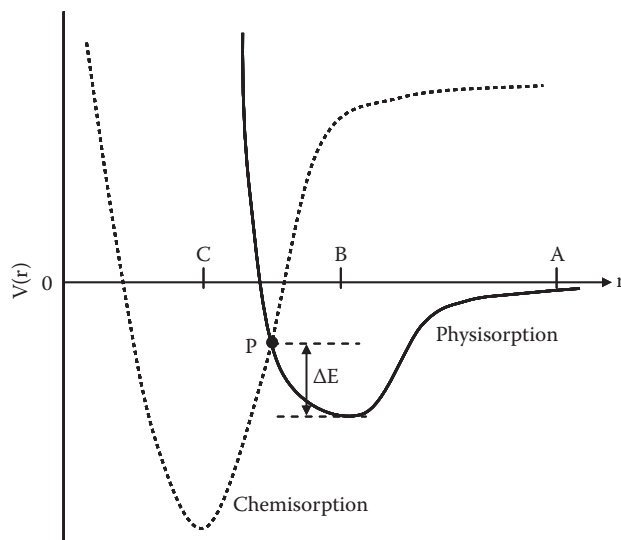
If adsorption is nonspecific and if only small amounts of heat are liberated, comparable to the heat of vaporization of the adsorbed material, the process is physisorption. The interaction between the adsorbent and adsorbate is much weaker ( $\sim 20 \text{ kJmol}^{-1}$ ) in a physisorbed film compared to a chemisorbed film. It should be noted that although physisorption is considered nonspecific, thermodynamic considerations are important when considering this process. As an example, we consider the physisorption of a long-chain alcohol, such as dodecanol, from an aqueous solution onto a solid hydrophobic surface. Let the hydrophobic surface be composed of an ODT monolayer chemisorbed to a gold surface. The adsorption process is driven by a strong tendency of the alcohol to avoid water (dodecanol is only sparingly soluble in water) and to associate with any accessible hydrophobic surface. Thus, the dodecanol adsorbs to the ODT SAM (Figure 3.8). The resulting dodecanol monolayer is closely packed with the OH headgroup oriented toward the water phase. Like the ODT SAM, the alkyl chains of dodecanol are closely associated with each other due to favorable van der Waals interactions, and form a solid-like SAM on the surface of the hydrophobic film. Unlike the ODT, the adsorption of dodecanol is reversible. The molecules in the film are in dynamic equilibrium with the dodecanol molecules in the aqueous bulk phase. Depleting the number of dodecanol molecules in the bulk phase reduces the packing density in the monolayer at the solid-aqueous interface.

Adsorption characteristics such as temperature, bulk concentration dependence, and reversibility are often used to distinguish chemisorption and physisorption. The energetic difference between the two adsorption processes is nicely



**Figure 3.8** The formation of a physisorbed monolayer of dodecanol on top of a SAM composed of octadecane thiol on a gold surface.

illustrated in the one-dimensional Lennard-Jones potential energy curve for the adsorption of a diatomic molecule on a planar surface.  $V(x) = 0$  at large distances corresponds to zero interaction between the substrate and adsorbate molecule at infinite separation. As the molecule approaches the surface (going from point A to B in Figure 3.9), there is a negative attractive potential between the surface and the adsorbate molecule. The potential reaches a minimum at distance B in the case of physisorption and at distance C in the case of chemisorption. These distances represent the corresponding equilibrium substrate-adsorbate bond lengths for the two processes. At distances less than B for physisorption (or less than C in the case of chemisorption), the attractive interaction lessens until  $V(x) = 0$ , below which repulsive interactions start to become



**Figure 3.9** A one-dimensional Lennard-Jones potential energy curve for the chemisorption and physisorption of a molecule on a planar surface. A, B, and C represent various distances between the surface and the molecule.

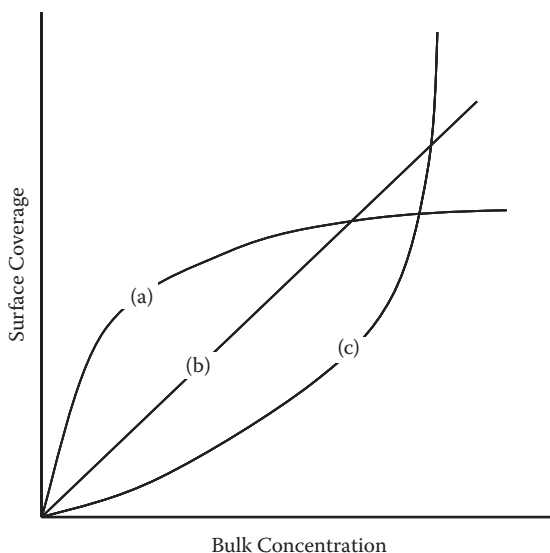
more important. The minimum potential (at distances B and C) for chemisorption is at a smaller substrate-adsorbate distance compared to physisorption because the former processes lead to a shorter bond distance between the substrate and adsorbate. Furthermore, the potential well for chemisorption is deeper than for physisorption due to a stronger substrate-adsorbate bond strength than that formed in physisorption, with the former process involving the breaking and making of covalent bonds. Sometimes a molecule can be trapped in a physisorbed state before being chemisorbed. In this case the physisorbed molecule is a precursor to chemisorption. The two potential energy curves shown in Figure 3.9 cross at a distance represented by point P. This is the point where the physisorbed precursor can “cross over” to the chemisorbed state.  $\Delta E$  represents the activation energy in going from the physisorbed state to the chemisorbed state.

The adsorption capacity of solid surfaces is determined from measurements of the mass (or moles) of material adsorbed and the area available for adsorption per unit mass of adsorbent. Usually the amount of solute adsorbed from solution per unit mass of adsorbent depends upon the solute bulk concentration up to its saturation point. Also, the amount adsorbed per

unit mass of adsorbent for a given solute bulk concentration decreases with increasing temperature. If the adsorption is physical, it is reversible and the solute leaves the surface of the adsorbent to reestablish equilibrium when the adsorbent is removed from a solution in equilibrium and placed in a solution of lower concentration. Such reversibility may not be shown if the process is chemisorption. The following section describes how adsorption capacity can be measured and how information, such as monolayer coverage and substrate surface area, can be extracted by measuring the adsorption capacity.

### 3.2.1 Simple Adsorption Isotherms

An adsorption isotherm is a plot of surface coverage (i.e., the fraction of the adsorbent surface covered by the adsorbate) versus the bulk concentration. The “isotherm” term means that these measurements are carried out at constant temperature. Figure 3.10 shows a few simple adsorption isotherms. The line that trails off to a constant surface coverage [Figure 3.10(a)] represents a situation in which adsorption stops because the surface is saturated by the adsorbate. This typically corresponds to monolayer coverage. In Figure 3.10(b), adsorption seems to continue indefinitely as a function of bulk concentration.



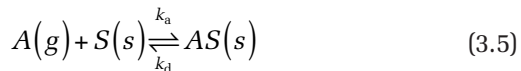
**Figure 3.10** Adsorption isotherms describing (a) Langmuir adsorption, (b) multilayer adsorption, and (c) cooperative adsorption.

Adsorption in this case does not stop after a saturated monolayer has formed; instead, adsorption continues and multilayers are formed on the surface. In Figure 3.10(c), adsorption seems to increase exponentially with bulk concentration. This situation occurs if the presence of the adsorbate at the surface promotes further adsorption due to favorable intermolecular interactions. This kind of adsorption is known as cooperative adsorption. It may or may not eventually level off at very high bulk concentrations.

If the bulk phase is a gas, then the adsorption isotherm will be a plot of surface coverage as a function of gas pressure at a constant temperature. Adsorption isotherms allow us to determine some important characteristics of the adsorption process such as equilibrium constants, the number of adsorption sites available for adsorption on the substrate, and the enthalpy of adsorption.

The Langmuir adsorption isotherm, presented by Langmuir in 1918, is the simplest adsorption isotherm. The Langmuir adsorption isotherm assumes that there are a finite number of adsorption sites on the solid substrate and that each site is taken up by a single adsorbate molecule. During the adsorption processes, these sites are taken up until we reach a point of saturation in which all sites have been taken up by adsorbate molecules. This point represents complete monolayer coverage, and further adsorption ceases. Furthermore, the Langmuir adsorption isotherm assumes that there are no interactions between the adsorbate molecules during the adsorption process, and that the enthalpy of adsorption is independent of surface coverage. Figure 3.10(a) represents a Langmuir adsorption profile.

By representing the solid surface by  $S(s)$  and the adsorbate molecules in a gas phase by  $A(g)$ , we can write the following equilibrium between the adsorbate and the substrate:



The constants  $k_a$  and  $k_d$  are the rate constants for the adsorption and desorption process, respectively. At equilibrium, the forward and reverse rates are the same, and so from basic kinetics,

$$k_a[A][S] = k_d[AS] \quad (3.6)$$

or

$$\frac{k_a}{k_d} = \frac{[AS]}{[A][S]} = K \quad (3.7)$$

where  $K$  is the equilibrium constant for the adsorption process. The adsorbate surface contains a certain concentration of sites available for adsorption. Let  $\beta$  be the concentration of sites per square meter and let  $\theta$  be the fraction of surface sites occupied by the adsorbate molecules. The concentration of occupied sites is therefore given by  $\beta\theta$ , and the concentration of free available adsorption sites is given by  $\beta - \theta\beta = (1-\theta)\beta$ . The rate of desorption ( $v_d$ ) is proportional to  $[AS]$ , which in turn is proportional to the number of occupied surface sites. Furthermore, the rate of adsorption ( $v_a$ ) is proportional to  $[A]/[S]$ , which in turn is proportional to the number of available sites and the number density of molecules in the bulk phase. Thus,

$$v_d = k_d\theta\beta \quad (3.8)$$

$$v_a = k_a(1-\theta)[A] \quad (3.9)$$

At equilibrium, these two rates are equal, so

$$k_d\theta\beta = k_a(1-\theta)[A] \quad (3.10)$$

Rearranging gives

$$\frac{1}{\theta} = 1 + \frac{k_d}{k_a[A]} = 1 + \frac{1}{K[A]} \quad (3.11)$$

For a gas phase adsorption process,  $[A]$  represents the concentration (number of molecules per unit volume) of the adsorbate gas. Assuming an ideal adsorbate gas, the concentration can be expressed as a pressure by use of the ideal gas equation,  $PV = nRT$ .

$$[A] = \frac{\text{number of molecules}}{V} = \frac{nN_A}{V} = \frac{nN_AP}{nRT} = \frac{N_AP}{RT} = \frac{P}{kT} \quad (3.12)$$



where  $P$  is the pressure of the adsorbate gas, and  $k$  ( $= R/N_A$ ) is the Boltzmann constant. Equation 3.11 can be rearranged to Equation 3.13.

$$\frac{1}{\theta} = 1 + \frac{kT}{KP} = 1 + \frac{1}{aP} \quad (3.13)$$

where  $a = K/kT$ . Equation 3.13 is the Langmuir adsorption equation. A plot of  $1/\theta$  versus  $1/[A]$  or  $1/P$  yields a straight line with a slope equal to the reciprocal of the equilibrium constant (or  $kT/K$ ).

The fraction  $\theta$  can be related to the following ratio:

$$\theta = \frac{V}{V_m} \quad (3.14)$$

In this equation,  $V$  is the volume of the gas adsorbed to the surface, and  $V_m$  represents the volume of gas corresponding to monolayer coverage. Complete monolayer coverage corresponds to  $\theta = 1$ , and  $V = V_m$ . Equation 3.14 can be substituted into Equation 3.13, giving Equation 3.15.

$$\frac{1}{V} = \frac{1}{aPV_m} + \frac{1}{V_m} \quad (3.15)$$

### Example 3.2 The Adsorption of Nitrogen onto a Mica Surface Follows the Langmuir Adsorption Model

The following data were collected at 273.15 K. Determine the value of  $a$ ,  $V_m$ , and the total number of surface sites. What would the concentration of surface sites be if the mica substrate was a square of length 2 cm?

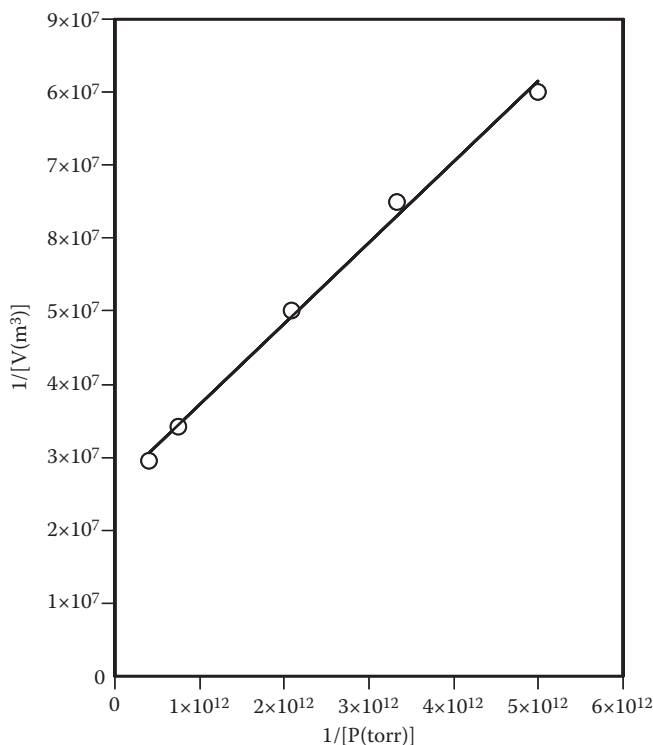
$P/10^{-12}$ torr	2.50	1.32	0.48	0.30	0.20
$V/10^{-8}$ m <sup>3</sup>	3.40	2.92	2.00	1.54	1.25

**Solution** According to Equation 3.15, a plot of  $1/V$  versus  $1/P$  will have a slope of  $1/aV_m$  and an intercept of  $1/V_m$ . Figure 3.11 shows this plot. A linear fit to the data yields the equation  $y = 1 \times 10^{-5}x + 3 \times 10^7$ .

The intercept on the y-axis is equal to  $3 \times 10^7 = 1/V_m$ , and so  $V_m = 3.3 \times 10^{-8}$  m<sup>3</sup>.

From the slope  $1 \times 10^{-5} = 1/aV_m$ , and so  $a = 3.0 \times 10^{12}$  torr<sup>-1</sup>.

To determine the total number of sites on the mica surface, we first need to determine the number of molecules corresponding to  $V_m$ .



**Figure 3.11** A plot of  $1/V$  versus  $1/P$  for the adsorption of  $\text{N}_2$  on mica at 273.15 K.

First, we know that under these conditions, 1 mol of gas occupies  $2.24 \times 10^{-2} \text{ m}^3$ .

The number of moles corresponding to  $V_m$  is

$$\frac{3.3 \times 10^{-8} \text{ m}^3}{2.24 \times 10^{-2} \text{ m}^3 \text{ mol}^{-1}} = 1.47 \times 10^{-6} \text{ mol}$$

The number of molecules is simply  $N_A$  times the number of moles.

$$(6.022 \times 10^{23} \text{ mol}^{-1})(1.47 \times 10^{-6} \text{ mol}) = 8.85 \times 10^{17} \text{ molecules}$$

Now each  $\text{N}_2$  molecule occupies a single site on the mica surface. Thus, there must be  $8.85 \times 10^{17}$  sites on the surface.

For a square of length 2 cm, the concentration of surface sites is simply the ratio of the total number of sites to the total surface area of the square.

$$\frac{8.85 \times 10^{17} \text{ molecules}}{(0.02 \text{ m})^2} = 2.21 \times 10^{21} \text{ m}^{-2}$$

### 3.2.2 Other Useful Adsorption Isotherms

The Langmuir isotherm makes a number of assumptions that may not apply to all adsorbates and adsorbents. For instance, strong intermolecular interactions may cause deviations from Langmuir adsorption behavior. Many systems display multilayer adsorption. The Brunauer Emmett Teller, or BET, isotherm is useful for modeling such behavior. The key assumptions made in the BET isotherms are that physisorption on a solid is infinite, that there are no inter-layer interactions in the multilayer film, and that each layer can be described by the Langmuir model. The BET isotherm takes the form shown in Equation 3.16.

$$\theta = \frac{V}{V_m} = \frac{cz}{(1-z)[1-(1-c)z]} \quad (3.16)$$

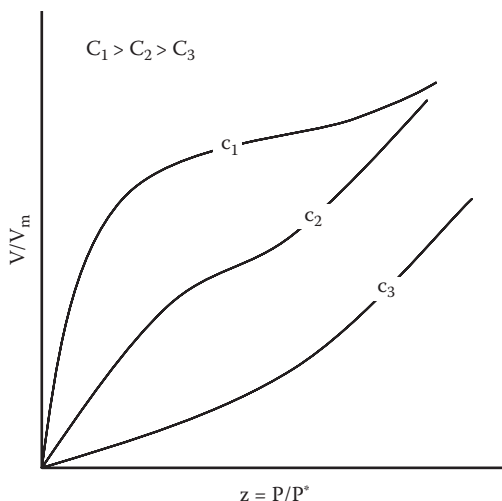
The term  $z = P/P^*$ , where  $P^*$  is the vapor pressure above the adsorbate before any adsorption occurs. The constant  $c$  is related to the enthalpy of adsorption of the first layer ( $\Delta H_1$ ) and the enthalpy of vaporization of the subsequent layers ( $\Delta H_{\text{vap}}$ ) (Equation 3.17).

$$c = \exp\left(\frac{\Delta H_1 - \Delta H_{\text{vap}}}{RT}\right) \quad (3.17)$$

Figure 3.12 shows how the form of the isotherm changes for different values of  $c$ . When  $c \gg 1$ , the isotherm simplifies to Equation 3.16 and is useful in describing surfactant adsorption (Section 3.3).

$$\frac{V}{V_m} = \frac{1}{1-z} \quad (3.18)$$

There are many useful isotherms, each one modeling a specific kind of adsorption. For example, the Temkin isotherm,  $\theta = c_1 \ln(c_2 P)$ , where  $c_1$  and  $c_2$  are constants, describes adsorption for systems in which  $\Delta H_1$  changes with pressure or concentration of the adsorbate. The Freundlich isotherm,  $\theta = c_1 P^{1/C_2}$ , describes adsorption involving strong intermolecular interactions. In Section 3.3.2, the Gibb's adsorption equation is introduced, which allows adsorption at the solid-aqueous interface to be modeled reasonably well.



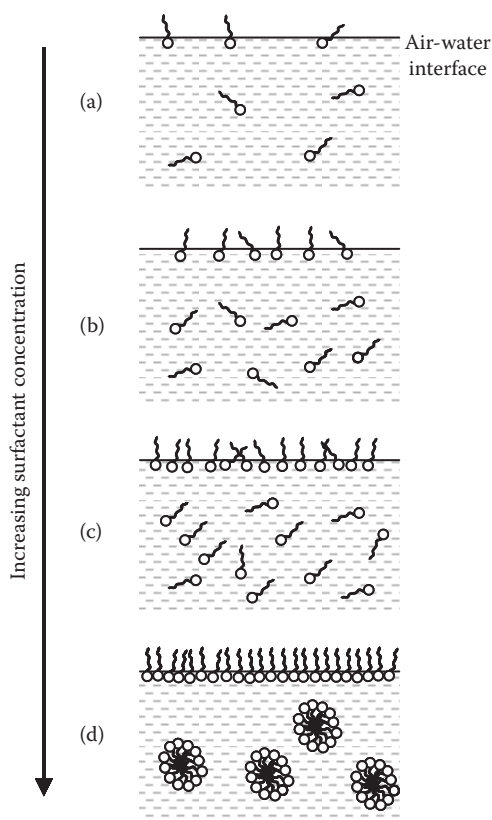
**Figure 3.12** The effect of changing the parameter  $c$  on the BET isotherm.

### 3.3 SURFACTANT CHEMISTRY

The word surfactant is an acronym for surface-active agent. Surfactants describe a class of molecules that have a tendency to adsorb to surfaces and interfaces and lower the interfacial tension. For instance, the air-water interface has a surface tension of 72.8 mN/m. The addition of 10% NaOH to the water increases this value to 78 mN/m. However, adding a typical surfactant to the water at relatively low concentrations (~mM) can lower the surface tension to 20 mN/m. Interestingly, the surface tension between an aqueous surfactant solution and an oil such as heptane can be as low as 1 mN/m.

To rationalize the effect surfactants have on surface tension, we need to consider the structure of these molecules. They are amphiphilic, which means that one part of the molecule is soluble in a specified fluid (the lyophilic part) and the other part of the molecule is insoluble (the lyophobic part). If the fluid is water, then the soluble part is called the hydrophobic part, and the insoluble part is the hydrophilic part. Thus, the structure of a surfactant has a region that is largely non-polar, typically hydrocarbon or fluorocarbon in nature, and another region that is polar, charged, and interacts strongly with water. These two regions are commonly referred to as the hydrophobic tail (or chain) and the hydrophilic head group, respectively. Figure 3.13 shows the structure of some common surfactants, including some naturally occurring ones such as phospholipids (DMPC).





**Figure 3.14** Surfactant molecules in the bulk aqueous phase are in equilibrium with a monolayer at the air-water interface. The population of the molecules increases with increasing bulk concentration (a)–(c). At the CMC (d) a saturated monolayer is formed and micelles are present in the bulk phase.

and a saturated monolayer is formed. The packing density of surfactant molecules within the monolayer will depend on the intermolecular interactions between neighboring head groups and the hydrophobic interactions between the tail groups. The formation of the monolayer represents a thermodynamically favorable situation: polar head groups are buried in the aqueous phase and the hydrophobic tails are avoiding contact with water. This scenario explains why surfactants lower the surface tension of the air-water interface. It is easy to see that surface tension decreases as the surfactant concentration increases, but then levels off to a constant value when a saturated monolayer is formed. In Chapter 4 we see how the surface tensions of surfactant solutions are measured.

### 3.3.1 Micelle and Microemulsion Formation

What happens if we continue to increase the surfactant concentration in an aqueous solution beyond that required to form a monolayer? The molecules are unable to adsorb to the air-water interface. Instead they self-assemble in the bulk phase into nanostructures or aggregates known as micelles. The shape and size of these micelles depends on the structure of the surfactant. For example, SDS forms spherical micelles; each micelle is 4 nm in diameter and is composed of about 60 molecules. The number of molecules making a micelle is known as the aggregation number.

The concentration at which a full monolayer is present at the interface and micelles begin to form in the bulk solution is known as the critical micelle concentration, or CMC. The process of micellization is in some ways similar to precipitation, but the precipitate itself has a very narrow size distribution and is stable and soluble in water. This property is because the micelle structure is such that the hydrophobic chains are aggregated in the core while the polar head groups form the exterior part of the structure [Figure 3.14(d)]. The delicate balance between inter-head group interactions and the geometry of the surfactant molecule result in a micelle with a particular shape, size, and aggregation number.

The CMC of a particular surfactant solution is very sensitive to impurities and other physical conditions. For example, increasing the temperature of an aqueous surfactant solution increases the CMC. Thermal agitation makes it more difficult for the molecules to self-assemble into micelles, so a higher concentration is required to reach the CMC. The addition of salt decreases the CMC of ionic surfactant solutions because the added ions screen the charged head groups of the surfactant, thus making it easier to form micelles. As the chain length of the surfactant increases, the CMC decreases due to the reduction in solubility of the surfactant. Table 3.2 lists the CMCs of some common surfactants.

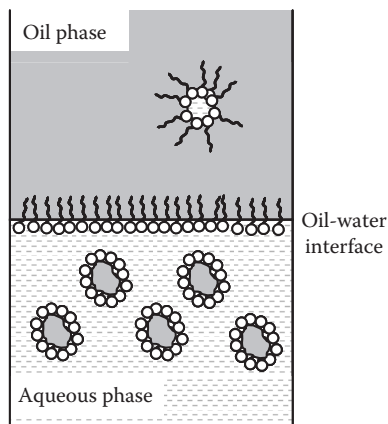
In a non-polar solvent, surfactant molecules can self-assemble to form “reverse micelles.” In this situation, the molecules are aggregated in such a way that the hydrophobic moieties form the exterior of the micelle and the polar regions of molecules form the core of the micelle. These micelles are thermodynamically stable in a non-polar solvent such as hexane.

Micelles in aqueous solution are capable of solubilizing small amounts of oil added to the aqueous surfactant solution

**Table 3.2** The CMC Values of Various Surfactants

Surfactant	CMC (mol dm <sup>-3</sup> )
Dodecylammonium chloride	$1.47 \times 10^{-2}$
Dodecyltrimethylammonium bromide	$1.56 \times 10^{-2}$
Decyltrimethylammonium bromide	$6.5 \times 10^{-2}$
Sodium dodecyl sulfate	$8.3 \times 10^{-3}$
Sodium tetradecyl sulfate	$2.1 \times 10^{-3}$
Sodium decyl sulfate	$3.3 \times 10^{-2}$
Sodium octyl sulfate	$1.33 \times 10^{-1}$
$\text{CH}_3(\text{CH}_2)_9(\text{OCH}_2\text{CH})_6\text{OH}$	$3 \times 10^{-4}$
$\text{CH}_3(\text{CH}_2)_9(\text{OCH}_2\text{CH})_9\text{OH}$	$1.3 \times 10^{-3}$
$\text{CH}_3(\text{CH}_2)_{11}(\text{OCH}_2\text{CH})_6\text{OH}$	$8.7 \times 10^{-5}$
$\text{CH}_3(\text{CH}_2)_7\text{C}_6\text{H}_4(\text{OCH}_2\text{CH})_6\text{OH}$	$2.05 \times 10^{-4}$

Source: Holmberg, K.; Jönsson, B.; Kronberg, B.; Lindman, B. 2003. *Surfactants and Polymers in Aqueous Solution*, 2nd ed., John Wiley & Sons, Chichester, West Sussex, England. With permission.



**Figure 3.15** When an aqueous surfactant phase above its CMC is brought in contact with a pure oil phase, the micelles are able to solubilize small oil droplets. The aqueous phase is a microemulsion. A monolayer of the surfactant likely populates the oil-water interface. The diagram also shows a “reverse” micelle in the oil phase containing a water droplet.

above its CMC. As a result the micelles swell with oil and increase in size (Figure 3.15). This swollen micellar phase is thermodynamically stable and is known as a microemulsion. In essence, tiny oil droplets are solubilized in water. Microemulsion phases are generally made by putting an aqueous surfactant phase in contact with an oil phase. The

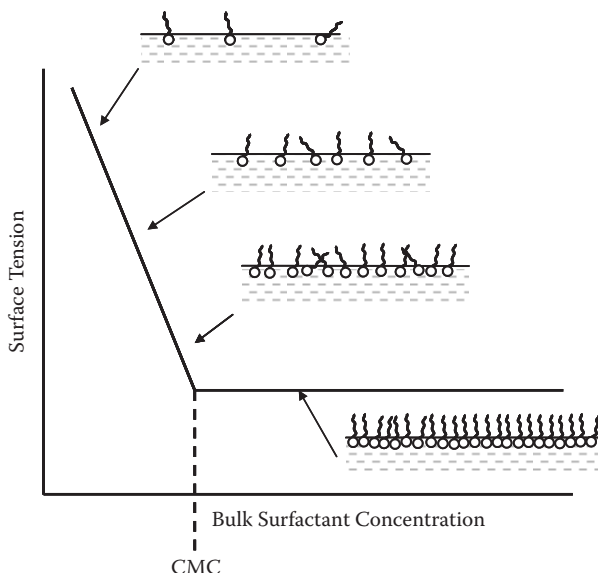


two phases don't mix completely, but a small equilibrium amount of the oil will enter the aqueous phase and turn it into a microemulsion. In fact, a complex equilibrium is established between the oil phase, a monolayer at the oil-water interface, and the aqueous micellar solution. It is also possible for reverse microemulsion aggregates to be present in the oil phase.

The amphiphilic nature of surfactants is the reason why these molecules self-assemble into nanostructures, both as micelles and microemulsions, and as monolayers. The ability to adsorb at interfaces and lower surface tension is the main reason why surfactants are used in almost all detergent and cleaning products. Microemulsions, with their ability to solubilize oil, have shown promise in tertiary oil recovery. Furthermore, manipulating the molecular structure while still preserving the amphiphilic nature allows the possibility of introducing a desirable functionality into the surfactant molecules and then exploiting its ability to self-assemble into novel materials.

### 3.3.2 The Determination of Surface Excess: The CMC and the Cross-Sectional Area per Molecule

Surface tension measurements of aqueous amphiphilic solutions allow one to determine the bulk concentration at which a saturated monolayer is reached. This concentration also corresponds to the CMC if a surfactant is used. Furthermore, one can use surface tension measurements to determine the cross-sectional area occupied by each molecule on the surface. Let's first consider a series of aqueous solutions of the cationic surfactant cetyl trimethylammonium bromide  $[(C_{16}H_{33})N(CH_3)_3Br]$ , or CTAB. CTAB is an effective antiseptic agent against bacteria and fungi and is used in hair conditioning products. After measuring the concentration of each solution, we can generate a surface tension versus concentration plot, as shown in Figure 3.16. The surface tension values drop from pure water ( $\sim 72 \text{ mN m}^{-1}$ ) to  $\sim 40 \text{ mN m}^{-1}$  when the bulk aqueous concentration of CTAB is 1 mM. The decrease in surface tension in this range is due to the increase in packing density as the bulk concentration is increased. Beyond 1 mM, the surface tension remains constant at  $40 \text{ mN m}^{-1}$ , which indicates that the packing density of CTAB molecules at the air-water interface is not increasing as the bulk concentration increases. At 1 mM and beyond, the surface has been saturated with CTAB molecules



**Figure 3.16** The drop in surface tension of air-water interface with increasing surfactant concentration. The intersection of the horizontal (constant surface tension) and the sloped lines corresponds to the CMC of the surfactant.

and a close-packed, highly oriented, monolayer is present on the surface.

What happens to the excess surfactant molecules in the bulk phase when the concentration exceeds that required to form a complete monolayer on the surface (1 mM in the case of CTAB)? They are unable to remain as individual molecules in solution and instead self-assemble into micellar nano-structures. This is shown in Figure 3.16, where a break point in a plot of the surface tension versus surfactant concentration marks the CMC. The CMC of an aqueous solution of CTAB is 1 mM.

Surface tension measurements also allow the determination of a quantity known as surface excess. Consider an aqueous solution containing surfactant molecules. Let's assume that we can determine the concentration of molecules in a two-dimensional plane somewhere in the bulk phase below the air-water interface. We can then compare this bulk phase concentration to the concentration of molecules on the actual surface of the solution. Since we are considering a surfactant solution, we expect the surface concentration of surfactant

molecules to be much higher than the bulk phase concentration. The difference between these two concentrations is known as the surface excess. The value tells us how much extra surfactant we have on the surface compared to a two-dimensional plane in the bulk phase. Surface excess is, to a good approximation, a measure of the concentration of molecules on the surface.

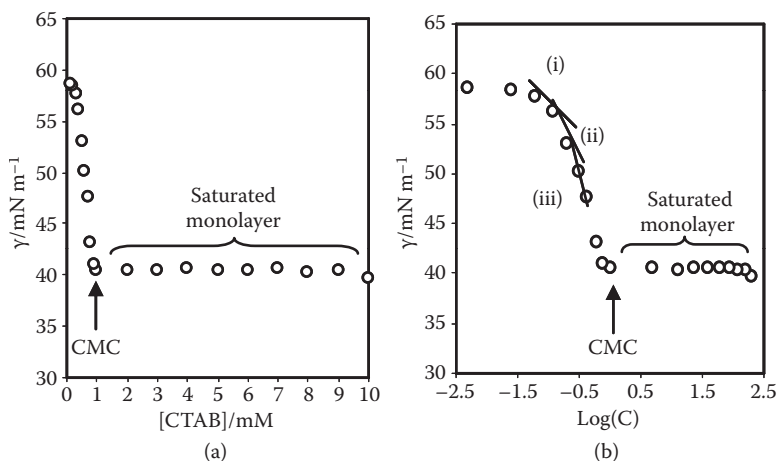
For non-ionic amphiphiles, the Gibb's adsorption equation relates surface tension ( $\gamma$ ) to surface excess ( $\Gamma$ ) at a given temperature ( $T$ ), and is given by

$$\Gamma = -\frac{1}{RT} \frac{d\gamma}{d \ln C} \quad (3.19)$$

where  $C$  is the surfactant concentration in solution below the CMC and  $R$  is the molar gas constant in  $\text{J K}^{-1} \text{mol}^{-1}$ . The surface excess is obtained from the slope of a plot of the surface tension versus the logarithm of the concentration. For ionic amphiphiles, a slightly different equation is used (Equation 3.20) because the presence of the counterion has to be taken into account and the surface as a whole must be considered electrically neutral.

$$\Gamma = -\frac{1}{2RT} \frac{d\gamma}{d \ln C} \quad (3.20)$$

Figure 3.17(b) shows a plot of the surface tension values versus the logarithm of the concentration for aqueous CTAB solutions. Below the CMC, the change in the slope reflects the different surface excess values at the various concentrations. At (i), the slope (or the value of  $d\gamma/d \ln C$ ) is relatively small, indicating a small surface excess. At the higher concentration indicated by (ii), the derivative  $d\gamma/d \ln C$  becomes larger, and according to Equation 3.19, the surface excess becomes larger. At higher concentrations like point (iii), the surface excess increases even more. The largest slope occurs at the CMC and the maximum possible surface excess is achieved, thus indicating the presence of a saturated monolayer at the aqueous-air interface. Equations 3.19 and 3.20 cannot be applied beyond the CMC.



**Figure 3.17** The surface tension of aqueous CTAB solutions versus (a) concentration and (b) logarithm of the concentration. The surface tension values were obtained at 20°C using the Wilhelmy plate method. Beyond the concentration corresponding to the CMC, the surface tension values become constant, indicating a saturated monolayer at the aqueous-air interface.

### Example 3.3 Calculating Surface Excess

Consider a concentrated aqueous solution of *n*-decanol. A plot of the surface tension versus the logarithm of the concentration gives a slope of  $-3.23 \text{ mN/m}$ . Use dimensional analyses to determine the units of surface excess and then calculate the surface excess. Using this value, determine the cross-sectional area per molecule of *n*-decanol at the air-water interface.

**Solution** Since the logarithm term in Equation 3.19 is unitless, the derivative  $d\gamma/\ln C$  has the same units as surface tension. The units of  $R$  are  $\text{J K}^{-1} \text{ mol}^{-1}$  and we know that  $1 \text{ N} = 1 \text{ J m}^{-1}$ . Thus, according to Equation 3.19, the units of  $\Gamma$  are

$$\frac{1}{(\text{J K}^{-1} \text{ mol}^{-1}) \cdot \text{K}} \text{ N m}^{-1} = \frac{1}{(\text{J K}^{-1} \text{ mol}^{-1}) \cdot \text{K}} \text{ J m}^{-1} \text{ m}^{-1} = \text{mol m}^{-2}$$

The surface excess of the *n*-decanol solution is

$$\begin{aligned} \Gamma &= -\frac{1}{RT} \frac{d\gamma}{d \ln C} = -\frac{1}{(8.314 \text{ J K}^{-1} \text{ mol}^{-1})(298 \text{ K})} (-3.23 \times 10^{-3} \text{ N m}^{-1}) \\ &= 1.30 \times 10^{-6} \text{ mol m}^{-2} \end{aligned}$$

Assuming the solution of *n*-decanol is close to saturation, we have close to a complete monolayer of the alcohol at the aqueous-air interface. The cross-sectional area per adsorbed molecule is inversely proportional to the adsorbed amount. If the surface excess is expressed in mol/m<sup>2</sup>, then the area per molecule,  $\sigma$ , is

$$\sigma(\text{m}^2/\text{molecule}) = \frac{1}{N_A \Gamma}$$

$\sigma$  is usually expressed in nm<sup>2</sup>/molecule. Since  $N_A = 6.023 \times 10^{23} \text{ mol}^{-1}$ , we have

$$\begin{aligned} \sigma(\text{nm}^2/\text{molecule}) &= \frac{(10^9 \frac{\text{nm}}{\text{m}})^2}{6.023 \times 10^{23} \frac{\text{molecules}}{\text{mol}} \cdot \Gamma} \\ &= \frac{1.6603 \times 10^{-6} \frac{\text{nm}^2 \cdot \text{mol}}{\text{m}^2 \cdot \text{molecules}}}{\Gamma} \\ &= \frac{1.6603 \times 10^{-6} \frac{\text{nm}^2 \cdot \text{mol}}{\text{m}^2 \cdot \text{molecules}}}{1.30 \times 10^{-6} \frac{\text{mol}}{\text{m}^2}} \\ &= 1.28 \text{ nm}^2/\text{molecule} \end{aligned}$$

Thus, each *n*-decanol molecule occupies an area of 1.28 nm<sup>2</sup> on the surface of water.

## REFERENCES AND RECOMMENDED READING

- Evans, D. F. and Wennerström, H. *The Colloidal Domain*, 2nd ed., 1999, Wiley-VCH, New York. pp. 99–153, 217–295. This is one of the finest books written on colloidal systems and includes a thorough discussion of surfactant chemistry, monolayers, and microemulsions. Also Chapters 3 and 5 are particularly relevant to the intermolecular interactions in nanomaterials.
- Holmberg, K., Jönsson, B., Kronberg, B., and Lindman, B. *Surfactants and Polymers in Aqueous Solution*, 2nd ed., 2003, John Wiley & Sons, Chichester, West Sussex, England. This is certainly one of the most thorough and comprehensible books on surfactant chemistry. It is very light on mathematical rigor and provides very thorough descriptions and examples.
- Rosen, M. J. *Surfactants and Interfacial Phenomena*, 3rd ed., 2004, John Wiley & Sons, New York. This book is an excellent source of data on many surfactants, including industrial synthetic surfactants, “green”

surfactants, and amphiphilic molecules relevant to living systems.

- Clint, J. H. *Surfactant Aggregation*, 1992, Blackie & Son. This book contains some excellent chapters on micelles and microemulsion. It is easy to follow and is highly recommended for anyone interested in surfactant science.
- Roach, P., Shirtcliffe, N. J., and Newton, M. I. "Progress in Superhydrophobic Surface Development," *Soft Matter*, 2008, 4, 224–240. This is an excellent review article on superhydrophobic surfaces and their relation to hard nanomaterials. The review contains many examples and references to methods used to construct superhydrophobic surfaces.
- Hartland, S., Ed. *Surface and Interfacial Tension: Measurement, Theory, and Applications*, 2004, Surfactant Science Series, Volume 119. Marcel Dekker, Santa Barbara, CA. This book provides a rigorous treatment of interfacial tension, film stability, and wetting phenomena. This book is recommended only for the serious student.
- Tóth, J., Ed. *Adsorption: Theory, Modeling, and Analysis*, 2002, Surfactant Science Series, Volume 119. Marcel Dekker, Santa Barbara, CA. This book contains many papers describing adsorption isotherms, including the adsorption behavior of biological molecules. The book is mathematically rigorous and is recommended only to those seriously interested in adsorption phenomena.

## END OF CHAPTER QUESTIONS

1. Calculate the work done when the surface of water increases by  $50 \text{ nm}^2$ .
2. Surfactant molecules on the surface of water, if sufficiently dilute, can be described as a two-dimensional gas phase. If there are no intermolecular interactions between the surfactant molecules, the gas can be considered ideal. Instead of the familiar  $PV = nRT$ , this ideal will obey a two-dimensional ideal gas equation  $\Pi\sigma = RT$ , where  $\Pi$  is the "surface pressure" due to the surfactant molecules, and  $\sigma$  is the surface area per molecule. Derive an expression for the reversible isothermal work due to expansion of this two-dimensional gas. What is the work done when the gas expands

and the area per molecule increases from 20 nm<sup>2</sup> to 40 nm<sup>2</sup>? Clue: The work can be determined by solving the following integral:

$$w = - \int \Pi d\sigma$$

3. The modified van der Waals equation is a more realistic equation describing a monolayer of lipid molecules at the air-water interface (see Question 2). This equation can be described as

$$\Pi = \frac{KT}{\sigma - \beta} - \frac{\alpha}{\sigma^2}$$

where the surface pressure ( $\Pi$ ) is a function of the independent variables temperature ( $T$ ) and the surface area per lipid molecule ( $\sigma$ ), i.e.,  $\Pi(T, \sigma)$ .  $K$ ,  $\alpha$ , and  $\beta$  are constants. Derive an expression describing the reversible isothermal work due to the expansion of a lipid monolayer obeying this equation of state.

4. Charcoal is an excellent adsorbate for organic molecules. The amount of dodecanol adsorbed on this material from a toluene solution was measured at room temperature. The following data gives the equilibrium amount adsorbed on charcoal and the corresponding equilibrium concentration of dodecanol in the bulk phase.

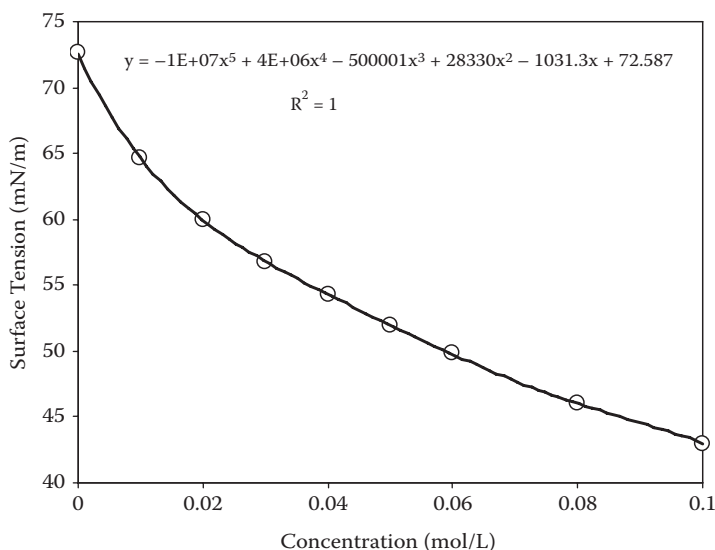
Bulk concentration (mol dm <sup>-3</sup> )	0.010	0.035	0.061	0.104	0.149
Amount adsorbed ( $\mu$ mol g <sup>-1</sup> )	24.0	50.3	70.0	81.2	90.8

- (a) By means of a graph, show that the data fit the Langmuir adsorption model. Calculate the area occupied by each adsorbed dodecanol molecule at saturation coverage. Take the adsorption area of the charcoal to be 100 m<sup>2</sup> g<sup>-1</sup>.
- (b) How would the adsorption isotherm change if the dodecanol was adsorbed from a slightly more polar solution compared to toluene?

5. This question concerns the application of Young's equation, which considers the equilibrium state of a drop of liquid on a solid surface in terms of the various surface tensions. Consider a drop of oil (an  $n$ -alkane) on the surface of solid polytetrafluoroethylene (PTFE). As expected, the oil spreads to some degree and forms a drop on the surface. The following oils were placed on the solid surface and the corresponding contact angles measured. The surface tension of each oil is also given.

Oil ( $n$ -alkane), $n$	$\cos\theta$	$\gamma$ (mN/m)
6	0.95	18
8	0.87	22
12	0.78	25
16	0.72	27

- Estimate the critical surface tension ( $\gamma_c$ ) corresponding to complete wetting. Show that its value is simply  $\gamma_{sv}$ . Do you think it is true that all liquids with a surface tension less than  $\gamma_c$  will spread on PTFE? If so, explain.
6. The Gibb's adsorption equation can be used to obtain values of surface excess at various concentrations. The following surface tensions were measured for aqueous solutions of  $n$ -pentanol at 20°C (see graph below). A polynomial function fit to the data is also shown along with the equation representing the best fit.



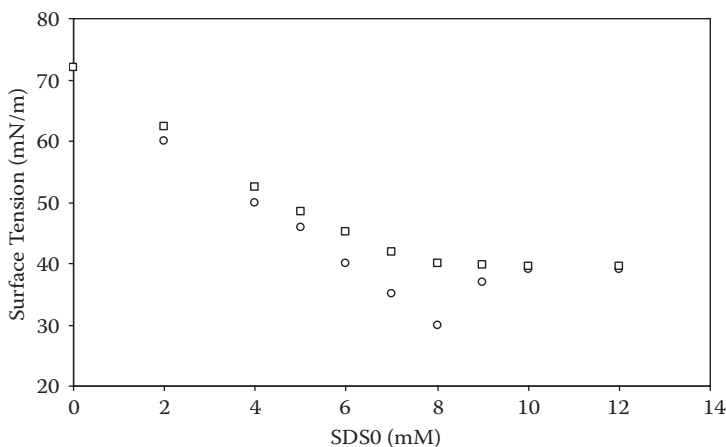


- (a) Calculate the surface excess concentrations and the average area occupied by each adsorbed molecule for bulk concentrations of 0.01, 0.02, and 0.04 mol dm<sup>-3</sup>.
- (b) Plot a surface pressure versus area per molecule ( $\Pi$ - $\sigma$ ) curve for the adsorbed *n*-pentanol monolayer and compare it with the corresponding curve for an ideal gaseous film. You should say something about how the molecules are interacting.  
Clue: The compressibility of the film is

$$Z = \frac{\Pi\sigma}{kT}$$

This should be equal to 1 at all surface pressures for an ideal gas.

7. Consider the following plot of surface tension versus SDS concentration. The squares represent the data from a purified sample of SDS (>99.99%, recrystallized thrice from ethyl acetate), and the circles show the data from a batch used as received from a popular vendor (>99.0%). The latter sample contains trace amounts of the corresponding alcohol (dodecanol).



- (a) Estimate the approximate CMC of pure SDS. Explain how this CMC is affected by (i) a decrease in temperature, (ii) an increase in ionic strength,

- (iii) an increase in pH, (iv) the addition of a *polycation*.
- (b) With the aid of diagrams explain the observation of the minimum in surface tension for the contaminated SDS sample. Explain clearly why the data differs from the purified sample.
8. In a recent paper in the *Journal of Chemical Education*, Bresler and Hagen reported an interesting laboratory project to investigate surfactant adsorption (Bresler, M. R.; Hagen, J. P. *J. Chem. Edu.* 2008, 85, 269-271). This problem will go through some of the steps these authors took to obtain an expression for the surface tension of a surfactant solution.
- (a) Rearrange the Gibb's adsorption equation (Equation 3.19) to the form

$$d\gamma = -\frac{RT\Gamma}{C}dC$$

- (b) The surface excess  $\Gamma$  can be assumed to follow the Langmuir adsorption isotherm, especially when considering a non-ionic surfactant. Express Equation 3.13 in the form

$$\theta = \frac{K[A]}{1 + K[A]}$$

and argue that

$$\sigma\Gamma = \frac{K_{\text{ad}}C}{1 + K_{\text{ad}}C}$$

where  $\sigma$  is the cross-sectional area of the surfactant molecules at the surface, and  $K_{\text{ad}}$  is the equilibrium constant for adsorption. What assumption has been made in saying that surfactant adsorption follows the Langmuir adsorption isotherm?

- (c) Solve

$$\sigma\Gamma = \frac{K_{\text{ad}}C}{1 + K_{\text{ad}}C}$$

for the surface excess, and then substitute the result into

$$d\gamma = -\frac{RT\Gamma}{C}dC$$

Integrate the resulting expression to obtain the Szyszkowski equation,

$$\gamma = \gamma_0 - \frac{RT}{\sigma} \ln(1 + K_{\text{ad}}C)$$

where  $\gamma_0$  is the surface tension of pure water.

9. The Szyszkowski equation in Question 8 can be used to obtain the standard Gibb's energy of adsorption ( $\Delta G_{\text{ad}}^{\circ}$ ). By plotting surface tension versus bulk concentration below the CMC and then curve fitting the data, one can obtain  $K_{\text{ad}}$ . This value is related to the standard Gibb's energy of adsorption by the equation  $\Delta G_{\text{ad}}^{\circ} = -RT \ln K_{\text{ad}}$ . The following data was collected for a non-ionic surfactant.

$C/\mu\text{molL}^{-1}$	0.0	2.5	5.2	8.0	13.0	17.5	21.0	31.5
$\gamma/\text{mNm}^{-1}$	72.8	53.6	49.7	45.2	42.6	40.8	40.7	40.8

- Plot  $\gamma$  versus  $\ln C$  and determine the CMC of the surfactant.
  - Plot  $\gamma$  versus  $C$  and determine the parameters  $\gamma$  and  $K_{\text{ad}}$  using nonlinear curve fitting.
  - Determine the standard Gibb's energy of adsorption.
  - The standard Gibb's energy of micellization is given by  $\Delta G_{\text{mic}}^{\circ} = RT \ln \text{CMC}$ . Determine  $\Delta G_{\text{mic}}^{\circ}$ .
10. Berberan-Santos commented on Bresler and Hagen's paper (Questions 8 and 9) in a letter published in the same journal (*J. Chem. Ed.* 2009, 86, 433). Berberan-Santos pointed out that the Szyszkowski equation can be reduced to the following equation if the surfactant concentration is sufficiently low:

$$\gamma = \left( \gamma_0 + \frac{\Delta G_{\text{ad}}^{\circ}}{\sigma} \right) - \frac{RT}{\sigma} \ln C$$

Derive this equation by considering a dilute surfactant solution. Use the data given in Question 9 to plot  $\gamma$  versus  $\ln C$ . From the linear plot obtain the values of  $\sigma$ ,  $\square G_{\text{ad}}^o$ , and  $K_{\text{ad}}$ . Berberan-Santos stated that nonlinear fitting is preferable in general (Question 9), but may not be mandatory. Discuss this in the context of the data presented in Question 9.

11. This question concerns the stability of spherical “nano-bubbles.” Consider a bubble, like that formed when a soapy film on a ring is blown. The bubble has an internal pressure  $P_1$ , and a radius  $r$ .  $P_o$  is the external pressure (e.g., the pressure of the surrounding air). At equilibrium, the bubble is stable and  $dG/dr = 0$ , where  $dr$  is the infinitesimal decrease in bubble radius. If  $P_1 > P_o$ , work must be done to ensure  $dr = 0$ . The change in Gibb’s energy due to the change in surface area is approximately equal to

$$dG = -8\pi r \, dr \, \gamma + \Delta P \, 4\pi r^2 \, dr$$

where  $\gamma$  is the surface tension of the bubble and  $\Delta P = P_1 - P_o$ . The first term in the above equation is a measure of change in Gibb’s energy due to surface tension, and the second term describes the mechanical work done against the pressure difference across the bubble surface.

Use the equilibrium condition  $dG/dr = 0$  to derive the Laplace equation:

$$\Delta P = \frac{2\gamma}{r}$$

The Laplace equation essentially gives the pressure inside a stable bubble.

- (a) The Young-Laplace equation

$$\Delta P = \gamma \left( \frac{1}{R_1} + \frac{1}{R_2} \right)$$

can be applied to nonspherical bubbles. In this expression,  $R_1$  and  $R_2$  represent the two principal radii of curvature. Show that the Young-Laplace

equation reduces to the Laplace equation for a spherical bubble.

- (b) Calculate the Laplace pressure ( $\Delta P$ ) in units of bars for bubbles of radius 1 nm, 2 nm, 10 nm, and 1000 nm. Comment on how  $\Delta P$  changes with the size of these “nano-bubbles.”
- (c) Using your understanding of Laplace pressure, explain why small boiling chips are often added to hot reaction mixtures.

## Four

---

# Characterization at the Nanoscale

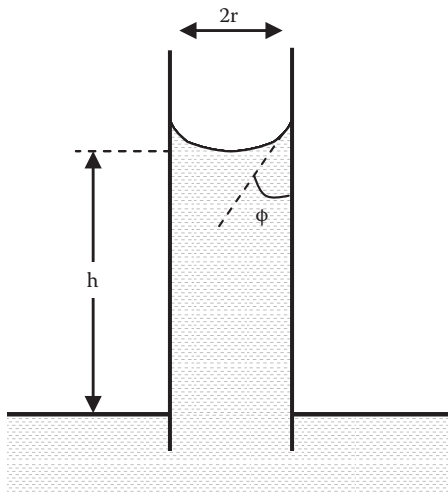
---

### CHAPTER OVERVIEW

This chapter provides a broad survey of some of the common techniques used to characterize nanomaterials. Novel techniques and methodologies to probe nanoscale dimensions are being developed continually, and many of these are based on variations of common tools and practices used in surface science and materials research laboratories. The aim of this chapter is to provide a modest level of background in a sufficient range of techniques to allow one to make rational choices in determining which method or combination of methods are best suited to characterize a particular feature of a nanomaterial. The methods introduced in this chapter range from traditional surface science tools, spectroscopic methods, and gravimetric techniques to more specialized characterization approaches such as nonlinear optical methods and interferometric techniques. The chapter focuses on the principles behind the techniques and the interpretation of data. Specific examples of the application of these methods to nanomaterials are found in Chapter 5.

### 4.1 SURFACE TENSIONOMETRY: THE SURFACE TENSIONMETER

There are a number of methods used to determine the surface tension of liquids. One important method is to study the movement of the fluid up through a capillary tube, and this is probably the most accurate method of determining surface tension values. The basic setup is shown in Figure 4.1. From this setup, the surface tension of a particular fluid of density  $\rho$  can be calculated directly using the height  $h$  to which the liquid rises through the narrow capillary of radius  $r$  as described by Equation 4.1.

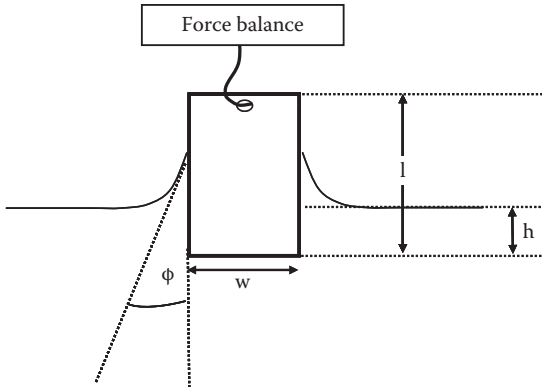


**Figure 4.1** A fluid moving through a narrow capillary tube. The distance  $h$  the fluid travels depends on the surface tension ( $\gamma$ ) and the contact angle ( $\phi$ ). The distance  $h$  also depends on the diameter  $2r$  of the capillary.

$$\gamma = \frac{rh\Delta\rho g}{2\cos\phi} \quad (4.1)$$

The term  $\Delta\rho$  represents the difference between the density of the liquid (e.g., bulk water) and the density of the vapor (usually air). The angle  $\phi$  is the contact angle the liquid makes against the capillary surface, and  $g$  is the acceleration due to gravity. For a narrow capillary the contact angle approaches zero, so the  $\cos\phi$  term in Equation 4.1 is generally set equal to 1. The surface tension then depends on the fluid density, the radius of the capillary, and the height the fluid travels up the tube. Other methods for determining surface tension include measuring the volume of a drop detached from a narrow tube, analyzing an image of a pendant drop, and observing a jet of liquid emerging from a nozzle of elliptical cross-section. A thorough treatment of these various methods can be found in volume 119 of the surfactant science series, *Surface and Interfacial Tension: Measurement, Theory, and Applications*.

The Wilhelmy plate method is another common technique used to determine surface tension values and is the one to which we confine our attention for the remainder of this section. The Wilhelmy plate method involves measuring the



**Figure 4.2** The Wilhelmy plate method. A platinum or paper plate of length  $l$  and width  $w$  is immersed into a fluid to a depth  $h$ . The forces acting on the plate are measured using a force balance.  $\phi$  indicates the contact angle the fluid makes against the plate. If a paper plate is used, this angle can be reduced to zero.

forces acting on a “plate,” usually a very thin piece of platinum or paper, at the liquid-air interface (Figure 4.2).

If a plate with dimensions  $l$  = length,  $w$  = width,  $t$  = thickness, and of density  $\rho_p$  is immersed to a depth  $h$  into a fluid of density  $\rho_L$ , then the forces acting on the plate are its weight, the upthrust on the submerged part of the plate due to buoyancy, and the surface tension of the liquid on the plate. The total force on the plate can be written as

$$F = (\rho_p l w t)g - (\rho_L h w t)g + 2(w+t)\gamma \cos \phi \quad (4.2)$$

where  $g$  is the acceleration due to gravity. In this expression,  $(\rho_p l w t)g$  is the weight of the plate,  $(\rho_L h w t)g$  is the buoyant upthrust, and the surface tension contribution to the force is  $2(w+t)\gamma \cos \phi$ , where  $\phi$  is the contact angle that the liquid makes on the plate as indicated in Figure 4.2.

#### Example 4.1 Measuring the Density from the Force Acting on a Plate

A clean plate of width  $w$  and thickness  $t$  is submerged into an aqueous surfactant solution. The solution has a surface tension value of  $\gamma$  mN/m. The force acting on the plate is measured as a function of immersion depth  $h$ . At  $h = 2$  cm, the force is  $F_1$ , and at a depth of  $h = 4$  cm, the force is  $F_2$ . Show how this information can be used to estimate the density of the surfactant solution.



*Solution* We can substitute the values of  $h$  and  $F$  into Equation 4.2 for the two cases and write,

$$\begin{aligned} \text{(a)} \quad & (\rho_p/wt)g - 0.02(\rho_L wt)g + 2(w+t) \gamma \cos\phi = F_1 \\ \text{(b)} \quad & (\rho_p/wt)g - 0.04(\rho_L wt)g + 2(w+t) \gamma \cos\phi = F_2 \end{aligned}$$

Subtracting these two equations gives  $0.02(\rho_L wt)g = F_1 - F_2$ .

Thus, by measuring the difference in force for a plate of a given width and thickness,  $\rho_L$  can be obtained.

The total force on the plate can be measured accurately by connecting the plate to a sensitive force balance. However, in order to accurately calculate the surface tension from the total force using Equation 4.2, we would need to know (also very accurately) the dimensions of the plate as well as the densities of both the plate and the liquid. Fortunately, the expression of the measured total force in Equation 4.2 can be simplified somewhat through careful manipulation of the procedure used to measure the total force on the plate, as follows.

First, before making any measurements and prior to submerging the plate into the liquid, the force balance is reduced to zero. This reduction eliminates the weight term from Equation 4.2. The plate is then submerged into the fluid and slowly raised until its lower edge is level with the liquid surface, thus eliminating the upthrust term because  $h = 0$ . In this position, Equation 4.2 is reduced to

$$F = 2(w+t)\gamma \cos\phi \quad (4.3)$$

Using Equation 4.3, therefore, we can determine the surface tension provided that we know the contact angle the liquid makes to the plate. This is often not necessary, however, because the expression can be further simplified by realizing that as we slowly pull the plate away from the liquid surface,  $\phi$  decreases and becomes zero (or  $\cos\phi = 1$ ) just before detachment. Thus, if we measure the force just before detachment of the plate from the liquid, we can determine the surface tension from the dimensions of the plate:

$$\gamma = \frac{F}{2(w+t)} \quad (4.4)$$

In order to obtain an accurate surface tension value, it is important to eliminate sources of contamination and work

at a constant temperature ( $\pm 0.5^\circ\text{C}$ ), since surface tension is temperature dependent. Furthermore, in order for Equation 4.4 to be used, a zero contact angle is required to exist between the plate and the liquid. This is usually achieved by using precisely cut paper plates, as the liquid can completely soak the paper, thus ensuring a zero contact angle. Platinum plates, even when cleaned thoroughly, can have their surfaces contaminated by the liquid during the first immersion, and their contact angle can vary during subsequent immersions. Flame cleaning the plate is usually recommended when working with platinum plates.

Surface tension experiments have been pivotal to our understanding of the properties of monolayers. The technique provides direct information on packing density and the area occupied by a molecule in a monolayer. Although the area per molecule gives a qualitative indication of the thickness of the film, surface tension cannot be used to give the absolute film thickness or the absolute mass of the monolayer. Fortunately, film thicknesses as low as a fraction of a nanometer can be measured using optical methods such as ellipsometry (Section 4.3) or dual beam polarization interferometry (Section 4.5). Gravimetric techniques also exist that can measure the mass of a monolayer on solid surfaces. The quartz crystal microbalance is a popular method used to determine the mass of a monolayer.

## 4.2 QUARTZ CRYSTAL MICROBALANCE

All gravimetric analyses rely on the determination of the mass of a material, or in our context, the mass of a nanofilm that has been deposited on a solid support. Since mass can usually be measured accurately, gravimetric analysis is generally a precise analytical method. However, when working with nanofilms, the mass in question may well be as low as a few nanograms per square centimeter, and traditional methods for weighing such samples are not possible. The next few sections discuss methods that allow the direct measurement of the mass and the thickness of a nanofilm (or, more often, the measurement of mass-related and thickness-related parameters). These methods are the quartz crystal microbalance with dissipation monitoring (QCM-D), ellipsometry, dual beam polarization interferometry (DPI), and surface plasmon resonance (SPR). These techniques are routinely used in many research laboratories and can measure the changes in mass and thickness during nanofilm growth.

First, let's focus on the quartz crystal microbalance (QCM). QCM is a powerful technique that can measure the mass of material as small as a few nanograms adsorbed to a surface. What makes this technique particularly appealing is that modern QCM instruments allow one to follow the mass deposition process as a function of time. In other words, the formation of a thin nanofilm, such as a model membrane, can be observed in real time. The quartz crystal microbalance is based on the piezoelectric characteristics of quartz, so in order to properly understand how QCM operates, let us first discuss the piezoelectric effect.

#### 4.2.1 The Piezoelectric Effect

Since early times it has been known that an electric field could be induced in certain types of crystals if they underwent a change in temperature (i.e., were heated or cooled). This phenomenon was named pyroelectricity. A Greek scientist, Theophrastus, first noticed the pyroelectric effect in 340 B.C. in experiments with tourmaline, a crystal silicate containing various metallic elements. In 1707, Johann Georg Schmidt showed that by heating or cooling tourmaline (as well as several other classes of crystals of proper geometries), a positive or negative electric field could be induced in the crystal.

In 1880, scientist-brothers Jacques and Pierre Curie discovered that they could also induce an electric field in certain classes of crystals, not by heating or cooling them, but by applying a mechanical stress. In addition, they realized that by changing the direction of the mechanical stress (i.e., a compression or expansion) they could control the sign of the electric field created. Their prior knowledge of pyroelectricity led them to this discovery, and they dubbed the phenomenon the piezoelectric effect, from the Greek *piezein*, meaning "to push." Today this is known as the *direct piezoelectric effect*—the ability of a crystal to produce an electric field in response to mechanical stress.

In subsequent years, the Curie brothers and other scientists began mapping out the crystalline structure requirements for piezoelectricity. They found that crystals possessing piezoelectric qualities have neither an absolute center nor a plane or axis of symmetry perpendicular to the axis of electric activity (the plane on which the electric field arises). There are currently 21 known classes of crystals that exhibit the piezoelectric effect, including quartz, cane sugar, and bone.

Another important advance in the understanding of piezoelectricity came with the discovery that a piezoelectric crystal

would be physically deformed (e.g., expand or contract) in response to an applied voltage. This effect was effectively the opposite of the direct piezoelectric effect and was subsequently named the *converse piezoelectric effect*. It was also discovered that the converse effect could be used to induce vibrations along a crystal. By applying a precise alternating current, the deformations of the crystal could be rapidly changed between expansion and contraction, causing the crystal to oscillate.

The piezoelectric effect is the foundation of a host of technologies. One of the original applications was a piezoelectric device used during the development of sonar in which piezoelectric quartz crystals were used as transducers to detect echoes returning from underwater objects. Piezoelectric devices are also commonly used in microphones—the air pressure fluctuations caused by a sound distort the piezoelectric device in the microphone, producing an electrical signal. The converse piezoelectric effect is also used for the extraordinarily fine and accurate adjustment of devices such as the needle in a scanning tunneling microscope.

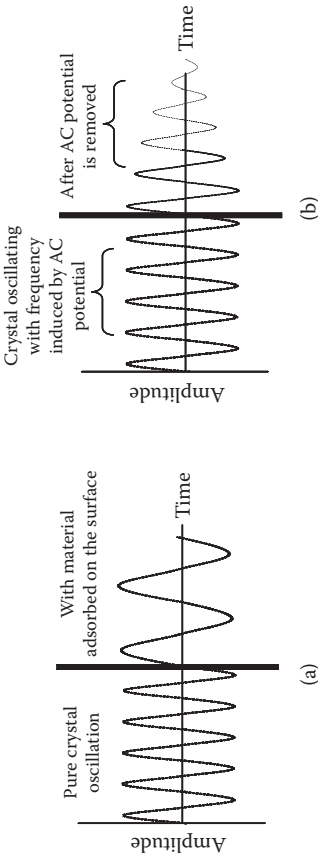
#### 4.2.2 QCM Principles

A quartz crystal microbalance operates on the principles of piezoelectricity. As we've discussed, a piezoelectric quartz crystal will oscillate at a specific resonant frequency in response to an AC voltage. (In fact, the resulting resonant frequency is so reliable that quartz crystals are used in many precision time-keeping devices). This resonant frequency is dependent upon the size and mass of the quartz crystal. If the mass of the quartz crystal is altered (such as when a molecule adsorbs to its surface), the resonant frequency at which the crystal is oscillating shifts slightly [Figure 4.3(a)]. If the adsorbed substance is not too thick or floppy, the change in mass on the crystal's surface is proportional to the shift in resonant frequency, as given by the Sauerbrey equation:

$$\Delta m = -\frac{C \cdot \Delta f}{n} \quad (4.5)$$

where  $m$  is mass,  $f$  is frequency,  $n$  is the overtone number of the crystal ( $n = 1, 3, 5, 7, \dots$ ), and  $C$  is a constant that depends upon the specific quartz crystal.

A quartz crystal microbalance detects minute changes in mass by applying an AC potential across a quartz crystal to induce its resonance frequency and then monitoring



**Figure 4.3** (a) The resonant frequency of a piezoelectric quartz crystal in a QCM system will oscillate more slowly when a material is adsorbed to its surface. Thus a decrease in the resonant frequency is observed during an increase in mass adsorbed to the surface. (b) Dissipation is a measure of how quickly the QCM crystal stops oscillating after the AC potential is removed. As such, it describes the relative thickness and rigidity of thin films adsorbed to the surface of the quartz crystal. The decaying oscillations of the crystal after the circuit is broken obey an exponentially decaying sinusoidal curve. The numerical data of the decay can be fitted to Equation 4.7, which allows for the calculation of a time constant  $\tau$ , which in turn can be used to calculate the dissipation of the crystal using Equation 4.8.

the changes in that resonant frequency that result from the adsorption of molecules to the crystal's surface. Since those resonant frequency shifts are proportional to mass under the conditions described above, and because even very small shifts in the resonant frequency can be detected with modern electrical equipment, QCM operates as a sensitive balance, or gravimetric device (hence the name microbalance). By monitoring the shifts in frequency as a function of time, QCM is able to monitor the formation of a thin nanofilm in real time. It should be noted, however, that the Sauerbrey relation shown in Equation 4.5 is often not exact when applied to surfaces at the solid-liquid interface because it was developed for oscillations in air and only applies to rigid masses attached to the crystal. It generally underestimates the mass adsorbed to the crystal surface under a liquid phase. Therefore, resonant frequency shifts detected using QCM should be considered a mass-related parameter, not an absolute measurement of mass adsorbed to a surface.

#### Example 4.2 What Are the Detection Limits of QCM?

A typical QCM-D instrument uses 5-MHz quartz crystals (meaning the crystal's fundamental resonant frequency is  $\sim 5$  MHz). These crystals have a Sauerbrey constant of  $C = 17.7 \text{ ng Hz}^{-1} \text{ cm}^{-2}$ . If the QCM-D instrument is capable of detecting changes in resonant frequency of  $\sim 0.1 \text{ Hz}$ , what is the detection limit of a typical QCM-D instrument at its fundamental resonant frequency ( $n = 1$ )?

*Solution* We utilize the Sauerbrey relation to find

$$\Delta m = -\frac{C \cdot \Delta f}{n} = -\frac{(17.7 \text{ ng} \cdot \text{Hz}^{-1} \cdot \text{cm}^{-2})(\sim \pm 0.1 \text{ Hz})}{1} = \sim \pm 2 \text{ ng} \cdot \text{cm}^{-2}$$

Therefore, the detection limit of a typical QCM-D instrument is  $\sim 2 \text{ ng cm}^{-2}$ . Indeed, the QCM is an incredibly sensitive balance!

#### 4.2.3 QCM and Dissipation (D)

The quartz crystal microbalance with dissipation monitoring (QCM-D), as compared with traditional QCM, offers additional information called dissipation about the materials adsorbed to the QCM crystal surface. Dissipation is a measure of the ability of the adsorbed material to release or dissipate the energy of the oscillating QCM-D crystal. As such, it provides insight into characteristics of the adsorbed film such as its density, thickness, and viscosity. Collectively, these properties are

sometimes referred to as viscoelastic properties. Dissipation is classically defined as

$$D = \frac{E_{\text{dissipated}}}{2\pi E_{\text{stored}}} \quad (4.6)$$

where  $E_{\text{dissipated}}$  is the energy lost during one oscillation cycle and  $E_{\text{stored}}$  is the energy stored in the crystal.

In practice, dissipation is determined by monitoring the time decay of the quartz crystal's oscillation when the AC potential is removed. The decay in the crystal's oscillations is an exponentially decaying sinusoidal of the form

$$A(t) = A_0 e^{-t/\tau} \sin(2\pi ft + \varphi) \quad (4.7)$$

where  $t$  is time,  $\tau$  is the decay constant,  $f$  is frequency, and  $\varphi$  is the phase angle. By numerically fitting the observed decay of the crystal's oscillations to Equation 4.7, the time constant  $\tau$  can be obtained, from which dissipation  $D$  can be calculated as

$$D = \frac{1}{\pi \cdot f \cdot \tau} \quad (4.8)$$

In other words, dissipation can be thought of as a measure of how quickly the crystal stops oscillating when the electrical circuit is broken [Figure 4.3(b)]. If the adsorbed thin film on the crystal's surface is thick and “floppy” (or not very rigid), then it is decoupled from the crystal's oscillations and efficiently dissipates the energy of the crystal. Consequently, the crystal stops oscillating quickly and a high value of dissipation is reported. Conversely, if the film is thin and rigid, it oscillates together with the crystal and does not dissipate the crystal's energy effectively. Hence, it takes a longer amount of time for the crystal to stop oscillating and a low dissipation value is reported. In this way, dissipation is a measure of the “floppiness” of the adsorbed film (or better, its lack of rigidity).

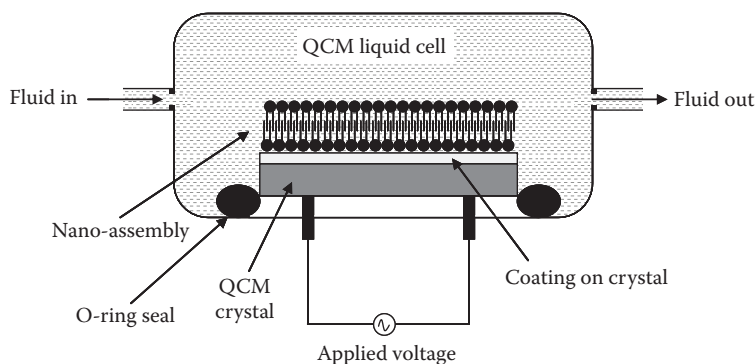
By using a suitable mathematical model, the dissipation and frequency shifts resulting from the adsorption of a nanofilm to a QCM-D crystal surface can be used to calculate the viscoelastic properties of the film. Viscoelasticity is the property of materials that exhibit both viscous and elastic

characteristics when undergoing some kind of deformation. Viscous materials resist flow when a stress is applied. Elastic materials when stretched very rapidly return to their original state once the stress is removed. Viscoelastic materials have elements of both of these properties. While the amount of valuable information obtained using such a mathematical model is impressive, caution should be observed in its use as the model inevitably requires the input of one known parameter (such as thickness, density, etc.) in order to calculate the other viscoelastic properties. If this required parameter is unknown or not known precisely, then the resulting calculated viscoelastic properties are unreliable. For these reasons, it is often helpful to perform a QCM-D experiment in conjunction with measurements taken from ellipsometry, SPR, or DPI.

#### 4.2.4 Modern QCM-D Setup

Traditional QCM was developed in the 1960s as a method to detect the adsorption of gas molecules to surfaces. With more recent advances, the capabilities of QCM have been extended to detect surface adsorption at the solid-liquid interface, and QCM-D has become a valuable tool in the characterization of thin films under solution.

A typical QCM-D setup involves the quartz crystal being mounted in a flow cell with electrodes mounted on either side of the crystal, as shown in Figure 4.4. The QCM-D crystal is generally prepared with an active sensor surface such as gold, hydroxyapatite,  $\text{SiO}_2$ , etc. The resonant frequency of the QCM



**Figure 4.4** Typical architecture of a QCM-D instrument. A QCM crystal is exposed to solution inside a flow cell. Changes to the crystal surface (such as film adsorption) are monitored via changes in the crystal's oscillations that have been induced by the applied voltage.



crystal is monitored as a function of time during the exposure of the crystal to a given solution under continuous flow conditions of  $\sim 0.100\text{--}0.300$  mL/min. If the solution contains materials that have some sort of affinity for the QCM crystal surface, they adsorb preferentially, increasing the mass of the quartz crystal and consequently producing a negative shift in the measured frequency. If the adsorbed material is sufficiently well behaved (i.e., is relatively thin and rigid), then the Sauerbrey relation can be applied, allowing us to calculate the mass of the adsorbed thin film.

QCM-D has a wide variety of applications. It is widely used in a biophysical context to detect interactions between biomacromolecules, such as between an enzyme and its substrate. It is also commonly used to examine the characteristics of polymer thin film formation as well as film behaviors under different conditions. Another application is determination of the effectiveness of different detergents to clean a given surface.

### 4.3 ELLIPSOMETRY

There are many approaches to determining the thickness of a nanofilm. Measuring the absorbance of light through a nanofilm is one way to determine its thickness. However, as will be shown, determining thickness from absorbance measurements requires knowledge of both molar absorptivity and the concentration of the material comprising the film. It may be challenging to obtain the concentration of the molecules in such a film.

Dragging a stylus over the surface of the substrate is another approach to determining film thickness. This method is the basis of the technique known as profilometry. A profilometer can be used to measure a film's thickness and roughness. In this method, a diamond stylus is moved vertically in contact with the surface of the nanofilm, then moved laterally across the surface for a specified distance and specified contact force. A profilometer can measure small surface variations in vertical stylus displacement as a function of position. A typical profilometer can measure small vertical features ranging in height from 10 nm to 50  $\mu\text{m}$ . In fact, most of the world's surface finish standards are written for contact profilometers. Unfortunately, the technique is "invasive" in that the stylus makes contact with the surface and may well damage the film. There are non-contact profilometers that use light as a way of measuring the height of surface features. However, non-contact profilometry

cannot measure the thickness of films on the order of  $10^{-10}$  m (angstroms).

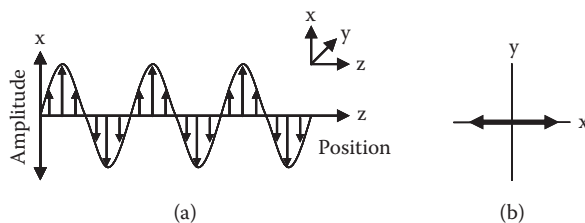
A technique known as ellipsometry has been used extensively to measure the film thickness as low as a few angstroms. Ellipsometry is a non-contact and nondestructive technique used for measuring the thickness and refractive index of a thin film on a solid surface. Thin films ranging in thickness from a few angstroms to  $\sim 1$   $\mu\text{m}$  can be measured accurately and quickly using this method. Ellipsometry analyzes the state of polarization of light reflecting from a surface and uses the laws of electromagnetism (specifically, Maxwell's equations) to resolve the thickness and refractive index of the nanofilm.

#### **4.3.1 Basic Principles of Electromagnetic Theory and Polarized Light**

Let's begin by describing the classical representation of light as an electromagnetic wave. Ellipsometry measures the change in the polarization of a light beam after its reflection from the solid surface of the sample being characterized. Although detailed coverage of the theory of ellipsometry is beyond the scope of this book, we include a brief overview of electromagnetic theory in order to better understand the basis of ellipsometry. The material also serves as background to other methods covered in this chapter such as DPI, SPR, and spectroscopic methods.

Light can be viewed as an oscillating electromagnetic field propagating through space. The oscillating field of light has two components: an electric field and a magnetic field that are mutually perpendicular to each other and also perpendicular to the direction of light propagation. Only the electric field component is considered here since the magnetic component does not interact appreciably with molecules. Thus, we consider light as an oscillating electric field whose amplitude and orientation can be represented by a line that we call the electric field vector. The orientation of the electric field vector at a given moment in time is defined as the "polarization axis" of the light. This model of light is shown in Figure 4.5, where the length and direction of the solid arrows indicate the strength of the field and its orientation, respectively.

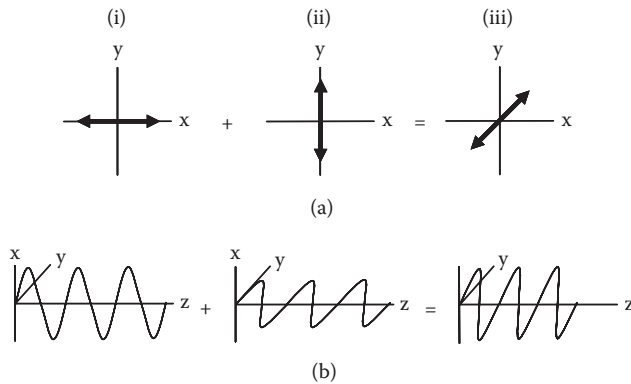
We now turn to the polarization of light. Usually, light emitted by most sources consists of photons whose electric fields are oriented in all directions that are perpendicular to the direction of propagation. This is unpolarized light. Conversely, linearly polarized light consists of photons whose electric fields are oriented in only one direction.



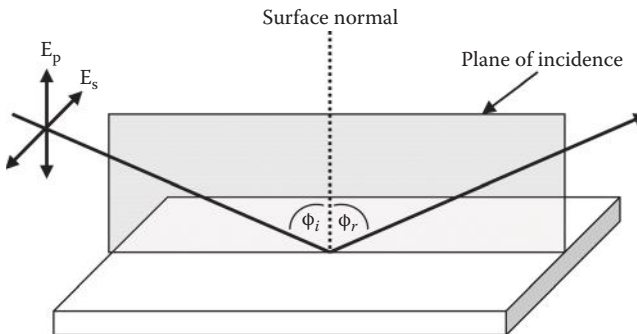
**Figure 4.5** Classical representation of light as an electric field. The solid arrows' vectors represent the field's orientation and intensity. The magnitude of the vectors oscillates in time, and their orientation defines the polarization axis. (a) is a wave representation of the light and (b) is a linear representation showing to which axis the electric field is confined.

In the simplest case of linearly polarized light, we can use a line to show the electric field oscillating along a single axis as shown in Figure 4.5(b). In fact, linearly polarized light, propagating along the  $z$ -axis can have its polarization axis oriented in any direction in the  $xy$ -plane. In other words, any source that produces linearly polarized light in the  $xy$ -plane can be thought of as being the linear combination of two vector components oriented along the  $x$ - and  $y$ -axes. The polarization axis of linearly polarized light is determined by the relative magnitudes of the two components. For linearly polarized light, the component light sources must have identical frequency and must also be in-phase with one another. By “in-phase” we mean that the minima and maxima of the electric field oscillations for each component must line up. A representation of linearly polarized light formed by a linear combination of two mutually perpendicular vector components is shown in Figure 4.6.

While we are on the topic of plane-polarized light, it is appropriate to introduce some common terminology when considering the reflection of light from a planar surface (Figure 4.7). For light incident upon this surface, we can define the plane of incidence as the plane that contains the surface normal and the light beam before and after the reflection (called the incident beam and reflected beam, respectively). If the electric field is polarized along the plane of incidence, we refer to the light as being p-polarized. Conversely, if the electric field is polarized along a plane perpendicular to the plane of incidence, we call the light s-polarized. A combination of s- and p-polarized electric fields can produce linearly polarized light, but only when there is a  $0^\circ$  phase difference between the two components. For example, if two s- and p-polarized light beams are



**Figure 4.6** (a) Linear and (b) wave representations of linearly polarized light of the same amplitude but various orientations. The polarizations are along (i) the  $x$ -axis and (ii) the  $y$ -axis. Vector addition of (i) and (ii) creates a linear polarization  $45^\circ$  from the  $x$ -axis (iii).



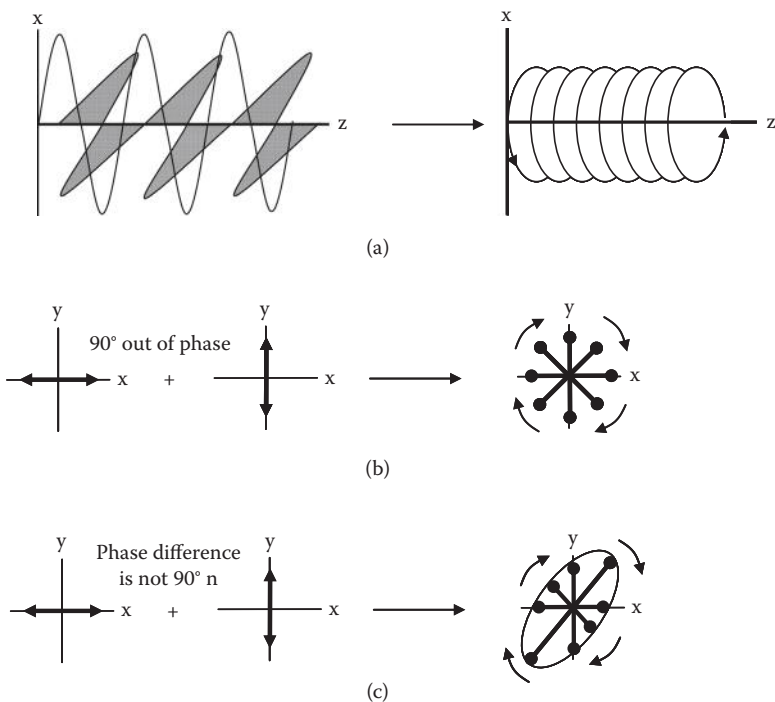
**Figure 4.7** Light reflection off a planar surface. The plane of incidence contains the incident beam, the surface normal, and the reflected beam. If the polarization of the light is along the plane of incidence, then it is called p-polarized light. If the polarization vector is perpendicular to the plane of incidence, then the light is called s-polarized light.

completely in-phase and if they each have electric field vectors of the same magnitude, then upon addition the resulting beam would be linearly polarized with a polarization axis that would be  $45^\circ$  with respect to the plane of incidence.

Let's now consider what would happen if the two electric field vector components making up a polarized light beam were not in-phase. First, let's talk about a special case—when the two components are  $90^\circ$  out of phase, or when one of the oscillations is a quarter of a wavelength ahead of the other.

When viewed along the direction of propagation, the resulting polarization has constant magnitude (the length of the vector line is the same), but its orientation changes in time such that the tip of the polarization vector traces a circular path in time. This kind of light is referred to as circularly polarized light and is depicted in Figure 4.8(a) and (b). Depending on the relative phases of the electric field components, the tip of the polarization vector may trace out either a left- or right-handed screw and the corresponding light is described as right- or left-circularly polarized.

If two out-of-phase light waves combine (and not with a phase difference of  $90^\circ n$  where  $n$  is an integer), then elliptically polarized light is formed. The linear combination of the two vector components results in a polarization vector that traces an ellipse in time when viewed along the axis of propagation. Elliptically polarized light is depicted in Figure 4.8(c).



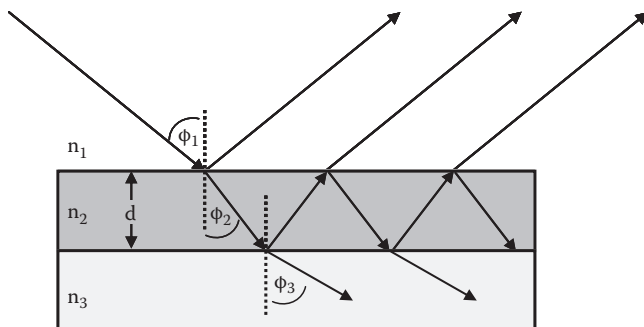
**Figure 4.8** (a) Two mutually perpendicular electric field components that are  $90^\circ$  out of phase will combine to generate circularly polarized light. (b) Vector addition of the two components leading to circularly polarized light. (c) Vector addition of the two components leading to elliptically polarized light.

As with circularly polarized light, elliptically polarized light can be left or right handed depending on the relative phases of the components. The “fatness” of the polarization ellipse (the ratio of the short axis of the ellipse to the long axis of the ellipse) is referred to as the ellipticity of the light. It should also be mentioned that elliptically polarized light can be produced by the linear combination of orthogonal components of circularly polarized light as well as linearly polarized light. In fact, linearly and circularly polarized light can be considered special cases of elliptically polarized light.

### 4.3.2 Basic Principles of Ellipsometry

Now that we’ve discussed the electromagnetic nature of light and polarization, let’s turn to our discussion of ellipsometry. Ellipsometry involves the reflection of light from a surface or the interface between two media. Consider a model reflective surface coated with a nanofilm of thickness  $d$  (Figure 4.9). For now we will limit our consideration to the interaction of an incoming light beam from a light source with the first surface (i.e., the interface between  $n_1$  and  $n_2$ ). In this model system, when light passes from one medium to a second medium, several phenomena occur at the interface. Some of the light is reflected from the surface and some enters the second medium.

When linearly polarized light reflects off a surface, there is a phase shift in both the parallel and perpendicular components (i.e., the s- and p-polarized components). As before, parallel and perpendicular are in relation to the plane of incidence of the incident light beam.



**Figure 4.9** The reflection, refraction, and transmission of light through a model multiple interface system. In our case, the layer of thickness  $d$  can be a nanofilm assembled on a solid support. The  $n_i$  represent the refractive index of each phase.

There may also be an amplitude difference between incident and reflected beams in both components. In fact, the phase shift and amplitude differences are usually not the same for both components. As a result, the reflected beam is elliptically polarized. The ellipticity depends on the optical properties of the substrate (reflecting surface) as well as the optical properties and thickness of any overlying films. One optical property of importance is the refractive index (RI), also denoted as  $n$ . The refractive index of a medium is defined as the ratio of the speed of light in a vacuum ( $3.0 \times 10^8$  m/s) to the speed of light in the medium.

It is useful to obtain a parameter that describes how much each of the s- and p-components of the incident light are reflected or transmitted. This information is given by the Fresnel reflection coefficient,  $r$ , which is the ratio of the amplitude of the reflected wave to the amplitude of the incident wave. For a single interface, the Fresnel coefficients for s- and p-polarized light are equal and are given by

$$r_{12}^p = \frac{n_2 \cos \phi_1 - n_1 \cos \phi_2}{n_2 \cos \phi_1 + n_1 \cos \phi_2} \quad (4.9)$$

$$r_{12}^s = \frac{n_1 \cos \phi_1 - n_2 \cos \phi_2}{n_1 \cos \phi_1 + n_2 \cos \phi_2} \quad (4.10)$$

In the two equations, the subscripts refer to medium 1 and medium 2 and the light transmitted through the interface is ignored.

So far, our discussion of ellipsometric principles has focused on the interaction of light with two media and a single interface. Now let's consider the rest of Figure 4.9, which shows a model of light interacting with multiple interfaces, in this case a reflective surface coated with a nanofilm of thickness  $d$ . The situation is only a bit more complex than for a single interface. The main source of complexity is that at each interface a light wave encounters, some of the light is reflected back from the interface and some is transmitted through it. Referring to Figure 4.9, we can realize that the result of this partial reflection/partial transmission is that some of the light that enters the thin nanofilm is internally reflected, "bouncing" between medium 1 and the reflective substrate surface. Furthermore, the intensity of the light inside the thin nanofilm eventually decays

as it releases some light back into medium 1 at each “bounce.” Each of these transmissions back into medium 1 is successively smaller and leads to a series of partial waves that combine to give a resultant total reflected wave. Therefore, our calculation of the amount of light reflected back into medium 1 from a system of multiple interfaces such as in Figure 4.9 must be a measurement of this total reflected wave and account for all of the small partial waves. This calculation is performed by modifying the Fresnel coefficients into total reflection coefficients  $R$  for multiple interfaces and these coefficients are given by

$$R^p = \frac{r_{12}^p + r_{23}^p e^{-i2\alpha}}{1 + r_{12}^p r_{23}^p e^{-i2\alpha}} \quad (4.11)$$

$$R^s = \frac{r_{12}^s + r_{23}^s e^{-i2\alpha}}{1 + r_{12}^s r_{23}^s e^{-i2\alpha}} \quad (4.12)$$

where

$$\alpha = 2\pi \left( \frac{d}{\lambda} \right) n_2 \cos \phi_2, \quad i = \sqrt{-1}$$

and  $\lambda$  is the wavelength of the incident light in a vacuum. The total reflection coefficients for each component (p or s) are the ratios of the reflected wave amplitude to the incident wave amplitude. We see that the equations for these total reflection coefficients each incorporate an exponential decay term,  $e^{-i2\alpha}$ , which accounts for the increasingly smaller partial waves that are produced by the partial reflection/partial transmission at the nanofilm/medium 1 interface. We also notice that if  $d = 0$  (or if there is no nanofilm on the surface) then Equations 4.11 and 4.12 reduce to Equations 4.9 and 4.10, as expected.

### 4.3.3 Obtaining the Thickness of Films: Optical Parameters $\Delta$ and $\Psi$

Section 4.3.2 covered the necessary background needed to understand how ellipsometry can be used to yield thickness values for thin nanofilms. At this point, we need to define two parameters ( $\delta_1$  and  $\delta_2$ ) that describe the change in phase as light is reflected off a surface. Let  $\delta_1$  be the phase difference



between the p-polarized component and the s-polarized component of the incident light. Let  $\delta_2$  be the phase difference between the p-polarized and the s-polarized component of the reflected light. We can now define one of the most important optical parameters used in ellipsometry, the parameter  $\Delta$ , as the phase difference between the p-polarized and s-polarized components of the incident light upon reflection. In other words,  $\Delta$  is the resulting change in the phase difference between the s and p waves as the light is reflected from the sample (Equation 4.13).

$$\Delta = \delta_1 - \delta_2 \quad (4.13)$$

The two components (p and s) making up the incident light each have a given amplitude (length of the electric field vector), and those amplitudes may also change upon reflection. These amplitude changes are given by the total reflection coefficients in Equations 4.11 and 4.12. The second fundamental ellipsometric optical parameter  $\Psi$  ( $\psi$ ) can be defined in terms of these coefficients and is given by the equation

$$\tan \psi = \frac{|R^p|}{|R^s|} \quad (4.14)$$

where  $\psi$  is defined as the angle whose tangent is the ratio of the magnitudes of the total reflection coefficients. We define the additional complex quantity  $\rho$  as the ratio of the total reflection coefficients, or

$$\rho = \frac{R^p}{R^s} \quad (4.15)$$

Using these three equations, we can now present the fundamental equation of ellipsometry:

$$\rho = \frac{R^p}{R^s} = \tan \psi e^{i\Delta} = \tan \psi (\cos \Delta + i \sin \Delta) \quad (4.16)$$

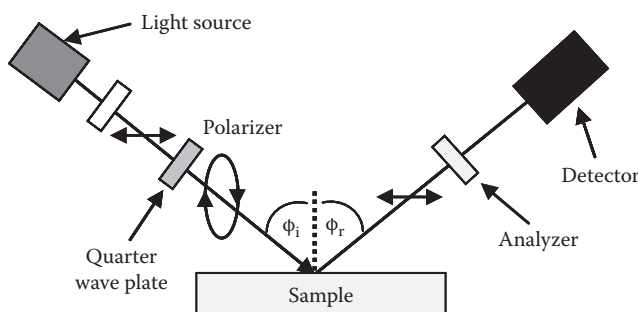
$\Delta$  and  $\psi$  are the experimental quantities measured by the ellipsometer and fitted to a computer model to yield refractive index and film thickness (which are embedded in the equations for  $R^p$  and  $R^s$ ).

Most ellipsometry experiments are performed on nanofilms at the solid-air interface. The underlying substrate is usually silicon. The determination of refractive index and film thickness for these samples is often done as follows. Suppose we have a substrate covered by a film as in Figure 4.9. Generally,  $\Delta$  and  $\psi$  are obtained for the bare silicon substrate and the instrument determines the refractive index information of silicon and the native  $\text{SiO}_2$  layer directly above it. This is done prior to any film deposition. A table of  $\Delta$  and  $\psi$  values as a function of film thickness (called a Del/Psi trajectory) is determined. The Del/Psi trajectory is obtained using a computer program separate from the ellipsometer, although some instruments have programs to compute Del/Psi trajectories integrated with the instrumentation. The film is then deposited on the substrate and  $\Delta$  and  $\psi$  are obtained for the substrate and film. The unknown thickness of the deposited film may then be obtained by comparison of the  $\Delta$  and  $\psi$  values with the calculated Del/Psi trajectory for the bare substrate. The optical constants of the thin film, such as the refractive index, must be input in order for the program to produce the trajectory. The user makes an educated guess at these values.

Although the ellipsometer accurately determines  $\Delta$  and  $\psi$ , these values are meaningless unless the program used to calculate the Del/Psi trajectory assumes the correct model. The model used is typically a two-layer model as shown in Figure 4.9 for a silicon dioxide/silicon wafer substrate.

#### 4.3.4 The Ellipsometer

The layout of a typical ellipsometer is illustrated in Figure 4.10. In short, a laser beam is appropriately polarized and then reflected off the substrate at an angle of  $70^\circ$ . The polarization



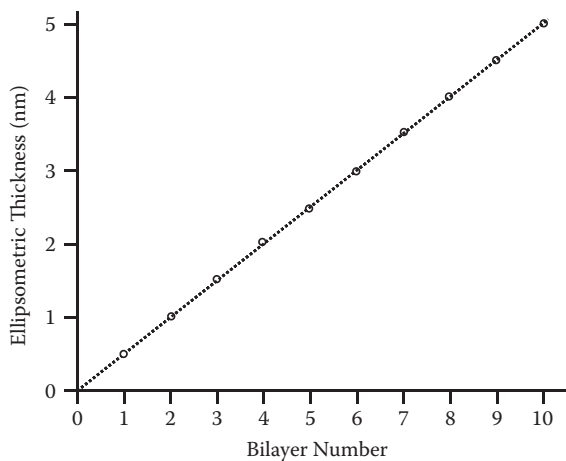
**Figure 4.10** Layout of a typical ellipsometer. The angle of incidence ( $\phi_i$ ) is typically  $70^\circ$ . The light source is a helium-neon laser, and the detector is a photomultiplier tube.

state of the reflected beam is measured using the appropriate optical equipment and its intensity is measured using a light-detecting photomultiplier tube.

The polarized light source is a helium-neon laser that emits a coherent beam of red light at a wavelength of 632.8 nm. Most common ellipsometers are usually equipped with a single-wavelength light source. However, because it is crucial that the nanofilm being studied does not absorb the incident light, spectroscopic ellipsometers have been developed that can produce light beams of several different wavelengths to avoid such problems.

The polarizer linearly polarizes the coherent unpolarized light beam from the helium-neon laser. The polarizer is a special filter that transmits light only if the polarization axis of the light lines up with that of the analyzer. If the light does not line up with the polarization axis of the polarizer, then the light is separated into its parallel and perpendicular components relative to the polarization axis and only the parallel component is transmitted. The linearly polarized beam then passes through another optical element, the quarter-wave plate. The quarter-wave plate is also commonly referred to as a retarder. Its purpose is to transform the linearly polarized beam into an elliptically polarized beam. The wave plate is an anisotropic material whose refractive index depends on the orientation of the propagating wave. P-polarized waves and s-polarized waves travel with different speeds through such a material. The thickness of the wave plate can be chosen to yield a beam whose components are exactly  $90^\circ$  out of phase with one another.

When linearly polarized light reflects off a surface, as previously explained, elliptically polarized light is typically produced. The rotating null ellipsometer produces an incident beam of varying ellipticities by varying the polarizer angle prior to the wave plate. When the beam reflects off the sample, the ellipticity of the beam changes. If the ellipticity is just right, then the change produced in the polarization of the beam by reflection from the sample produces a linearly polarized beam. The analyzer, which is identical to the polarizer, is then rotated until the polarization axis of the analyzer is perpendicular to the polarization axis of the reflected beam. The polarizer and analyzer angles are rotated sequentially until a null is located. These angles are used to determine the optical parameters  $\Delta$  and  $\psi$  for the sample. In practice, the instrument completes this computation and the values for  $\Delta$  and  $\psi$  are output.



**Figure 4.11** Thickness data as determined by ellipsometry for the layer-by-layer construction of a typical polyelectrolyte nanoassembly. The bilayer represents a layer of polycation complexed with a layer of polyanion.

A beautiful illustration of ellipsometry involves characterizing the sequential deposition of a polycation and a polyanion on a silicon surface. Polyelectrolyte nanoassemblies (the protocols and applications) are discussed in Chapter 5. Essentially, a clean silicon substrate is immersed into a solution of a polycation for about 5 minutes. It is then removed, washed, dried, and then immersed into a solution of a polyanion. The adsorption of the polyelectrolytes is driven by electrostatic attraction for each other, and the resulting assembly can be described as a polyelectrolyte bilayer. The procedure can be repeated to form many layers.

After the construction of each layer, ellipsometry can be used to determine the film thickness. The thickness can be determined to about 0.2 nm. Usually some effort has to be made to ensure that the thickness is determined from the same area on the substrate. Figure 4.11 shows how the thickness of a layer-by-layer polyelectrolyte nanoassembly changes with each successive layer.

#### 4.4 SURFACE PLASMON RESONANCE

Other techniques aside from ellipsometry exist to determine the refractive index and thickness of a thin nanofilm. A surface plasmon resonance (SPR) sensor is another optical method that is employed to detect changes in thickness and refractive index of very thin organic films adsorbed to a metallic

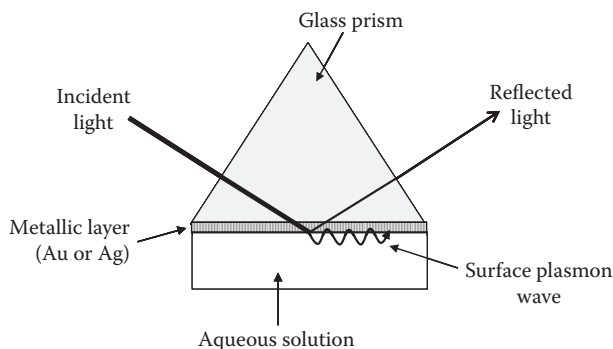
surface. The SPR is often used to detect interactions between molecules, and it has emerged as perhaps the most widespread “surface method” for detecting and quantifying interactions between biological macromolecules at the nanoscale.

#### 4.4.1 Principles of SPR

When incident light strikes the interface between a substance with a high index of refraction and another substance with a lower index of refraction, the light is completely reflected as long as the angle of the incident light is greater than that of the critical angle (see section on Evanescent Waves in the Appendix). In fact, the critical angle  $\theta_{\text{critical}}$  for the interface between any two substances can be calculated by applying Snell's law:

$$\theta_{\text{critical}} = \arcsin\left(\frac{n_2}{n_1}\right) \quad (4.17)$$

where  $n_1$  and  $n_2$  are the indices of refraction of the more dense and less dense substances, respectively ( $n_1 > n_2$ ). This phenomenon of complete reflection of the incident light is termed total internal reflection. Total internal reflection is normally observed when visible light is shone upon the interface between a glass prism ( $n = \sim 1.5$ ) and water ( $n = \sim 1.3$ ) at  $\theta_{\text{incidence}} > \theta_{\text{critical}}$ . However, if the surface of the prism facing the aqueous solution is coated with a thin layer of silver or gold as shown in Figure 4.12, then total internal reflection



**Figure 4.12** The generation of a surface plasmon wave. When polarized light is shone at the correct angle through a glass prism onto a metal-water interface, then a surface plasmon wave is generated that propagates along the interface. This results in a reduction of intensity of the reflected light.

is not always observed. This loss of total internal reflection occurs because some of the incident light is “channeled” into the metal-water interface where it generates oscillating waves of surface charge density that move along that interface. These oscillating waves of surface charge density are called surface plasma waves or surface plasmons, and the phenomenon of their creation serves as the basis for SPR sensors.

The creation of these surface plasmons is angle dependent, meaning that there exists an angle of the incident light (greater than  $\theta_{\text{critical}}$ ) at which the generation of the surface plasmons reaches a maximum. This angle is defined as the surface plasmon resonance angle  $\theta_{\text{spr}}$  (or SPR angle). The SPR angle therefore represents the greatest “channeling” of the incident light into the metal-water interface and consequently the greatest reduction in the intensity of the reflected light. So by measuring the reduction in reflectance as a function of the angle of incidence, the SPR angle (or angle at which a global minimum in reflectance is observed) can be determined.

Of more practical use, it turns out that the electric field associated with the surface plasmons is not completely contained in the metal-water interface, but stretches slightly into each medium, decaying exponentially as it extends away from the interface. Because of this penetration into each medium, any changes to the refractive index near the metallic surface alter the properties of the electric field, which in turn alters the SPR effect. Therefore, the SPR angle is sensitive to even very small changes in the index of refraction near the metal-water interface, such as might occur when a protein or polymer film is adsorbed to the metallic surface. In fact, for a given wavelength  $\lambda$  of light, the change in  $\theta_{\text{spr}}$  is related to the change in refractive index at the surface  $\Delta n_{\text{surface}}$  and the change in the thickness of the thin film  $\Delta d$  by

$$\Delta\theta_{\text{spr}}(\lambda) = c_1\Delta n_{\text{surface}} + c_2\Delta d \quad (4.18)$$

where  $c_1$  and  $c_2$  are constants.

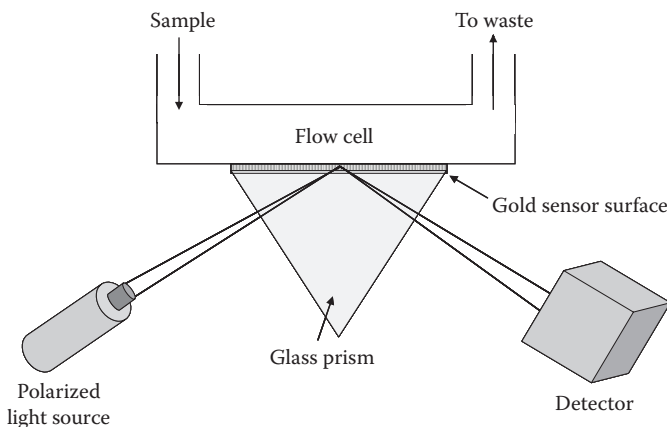
An SPR sensor typically operates by detecting changes in the SPR angle during the adsorption of molecules to the metal surface. Because  $\Delta\theta_{\text{spr}}$  is a function of  $\Delta n_{\text{surface}}$  and  $\Delta d$ , which in turn are a result of molecular interactions at the surface, an SPR sensor serves as a sensitive method to detect binding events of molecules to a surface. It should also be noted that

$\Delta\theta_{\text{spr}}$  is typically reported in resonance units (RU), where 1000 RU corresponds to a change of  $0.1^\circ$ . For common analytes such as proteins, the correlation between  $\Delta\theta_{\text{spr}}$  and the amount of substance adsorbed to the surface has been determined and a 1000-RU shift is approximately equivalent to an adsorption of  $0.1 \text{ ng/mm}^2$  for most proteins.

It should be noted from Equation 4.18 that the SPR angle is a function of both the thickness  $d$  and the index of refraction  $n$  of the thin film, meaning that neither parameter can be extracted by itself without making certain assumptions (i.e., a change in thickness could be extracted if it were assumed that the refractive index remained constant). For this reason, the change in SPR angle is said to measure the *effective refractive index* or a *thickness- and RI-related* parameter. In order to resolve thickness and absolute refractive index simultaneously, another method must be used, such as dual polarization interferometry.

#### 4.4.2 SPR Instrument Setup

A schematic diagram of a common SPR sensor (such as the prevalent Biacore instrument) is shown in Figure 4.13. The primary component is a prism that has been coated with a gold sensor surface so that surface plasmons are generated when



**Figure 4.13** Typical setup of an SPR instrument. A polarized light source is used to excite the SPR effect at the gold-water interface. The detector monitors the changes in the intensity of the reflected light as a function of angle of incidence in order to detect changes in the SPR angle that may result from the adsorption of a thin nanofilm to the gold sensor surface. A flow cell allows for the easy exposure of the gold sensor surface to the desired sample solution.

polarized light of a particular wavelength is shone upon it at the SPR angle. This gold sensor surface serves as the bottom of a flow cell through which solutions of various analytes can be passed. If the analytes interact with the sensor surface (or with molecules that have been pre-adsorbed to the surface), then a change in the SPR angle is observed by the detector and recorded by the computer software. The flow cell itself is typically very small (20–60 nL) and common flow rates are on the order of 1–100  $\mu\text{L}/\text{min}$ .

While SPR sensors can be employed for a variety of purposes that involve the characterization of ultra-thin films, they are primarily used as biosensors—to detect interactions between proteins and substrates, between strands of DNA, and between drug molecules and protein targets. SPR can also be used to monitor the formation of self-assembled monolayers on gold surfaces, particularly the formation of alkyl-thiol monolayers.

The SPR effect is not limited to planar surfaces, but can also be observed using gold or silver nanoparticles in solution. In this case, the nanoparticles have a wavelength-dependent absorbance that is a result of the excitation of surface plasmons within the nanoparticles, and this wavelength is measured rather than an angle-dependent response as with planar SPR (which would be difficult to measure for a nanoparticle). The wavelength at which the maximum absorbance is observed is a function of the particle size and particle shape. For example, spherical gold nanoparticles with a diameter of  $\sim 13$  nm exhibit a maximum absorbance at 520 nm.

The wavelength at which the SPR effect is maximized also depends on the dielectric environment (i.e., the refractive index) near the nanoparticle surface and on the distance between neighboring nanoparticles. These two properties have allowed gold nanoparticles to be used as sensitive biosensors. For example, short pieces of DNA called oligonucleotides have been covalently attached to gold nanoparticles through thiol chemistry, and when complementary DNA is introduced into the solution, the oligonucleotide-coated nanoparticles aggregate as they hybridize to the strands in solution. This aggregation causes the distance between neighboring nanoparticles to decrease and a blue-shift in the wavelength of maximum absorbance is observed. In this manner, gold nanoparticles can be used to determine the complementarity of two DNA sequences. Similar methods have also been developed to examine the changes in refractive index that occur on the surface of a gold nanoparticle when something



adsorbs to its surface—for example, when an antigen binds to a surface-bound antibody.

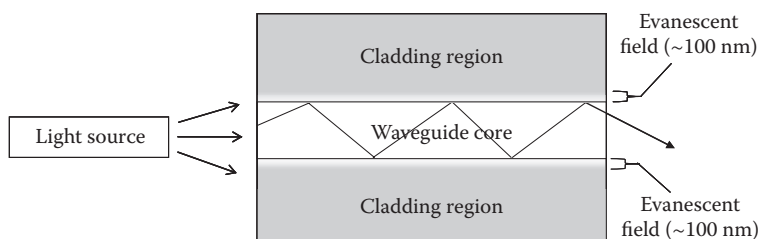
## 4.5 DUAL POLARIZATION INTERFEROMETRY

Unlike other surface characterization techniques such as QCM-D and SPR whose measurements offer only indirect estimations of the mass, density, or thickness of a thin film, dual polarized interferometry (DPI) is an optical technique that actually provides an exact method to measure these parameters at the solid-liquid interface. Since DPI provides two independent measurements of the effective refractive index of a thin nanofilm, it is able to simultaneously determine the mass, density, and thickness of that nanofilm. A variety of other waveguide spectroscopic techniques exist that are similar to DPI, such as optical waveguide lightmode spectroscopy (OWLS) and coupled plasmon waveguide resonance (CPWR) spectroscopy, but only DPI is discussed here.

### 4.5.1 Waveguide Basics

As may be recalled from our discussion of the critical angle of reflection, if light is shone on the boundary between two substances, the light is totally reflected within one substance as long as it is shone at an angle greater than the critical angle ( $\theta_{\text{critical}}$ ) and if the refractive index of the second substance is less than that of the first (as shown in Equation A.2 in the Appendix). Such a phenomenon is termed total internal reflection.

With that in mind, suppose we were to construct a layer of a substance with one refractive index and sandwich it between two layers of a substance with a lower refractive index. In such a setup, we would expect that light shone at an angle greater than  $\theta_{\text{critical}}$  could undergo total internal reflection inside the middle layer, alternately reflecting off the top and bottom layers as it passed through. Indeed, such behavior is observed and the setup is called a waveguide, with the middle layer termed the waveguide core and the top and bottom layers called the cladding regions, as shown in Figure 4.14. When light is shone upon the edge of a waveguide, total internal reflection can occur and the light passes through the core and emerges from the other side of the waveguide. As a side note, a similar application of total internal reflection is responsible for the ability of fiber optic cables to transmit light over long distances (the interior of the fiber optic cable is of a higher refractive index than the surrounding material).



**Figure 4.14** The setup of a basic waveguide. A waveguide is typically composed of a waveguide core sandwiched between two cladding regions. Light is transmitted through the waveguide core via total internal reflection. At each surface between the core and the cladding regions, an evanescent field is generated that extends  $\sim 100$  nm into the cladding region.

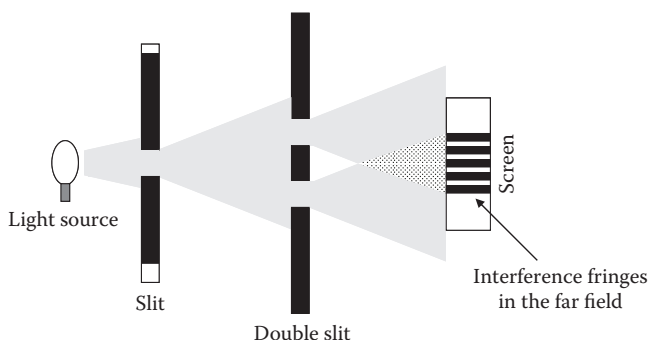
When light undergoes total internal reflection within the waveguide, it turns out that the light is not completely contained in the core, but it extends slightly into the cladding regions. This extension is called the evanescent field (see Equation A.1 in the Appendix). The evanescent field decays exponentially, moving away from the surface of the waveguide core, and at a distance of  $\sim 100$  nm into the cladding region its presence is negligible. Because of the interaction that occurs between the evanescent field and the cladding region, the speed at which the light propagates through the core of the waveguide depends slightly on the refractive index of the area of the cladding regions near the core's surface. Thus, light traveling through a waveguide can be a sensitive measure of any changes to the refractive index that occur within the first hundred nanometers of the core's surface, such as when a thin nanofilm is deposited upon the surface of the waveguide core. The problem then becomes how to accurately measure the speed at which the light travels through the waveguide (or in other words, how to measure the refractive index of the film deposited on the core's surface). To solve this issue, we turn to interferometry.

#### 4.5.2 Waveguide Interferometry and the Effective Refractive Index

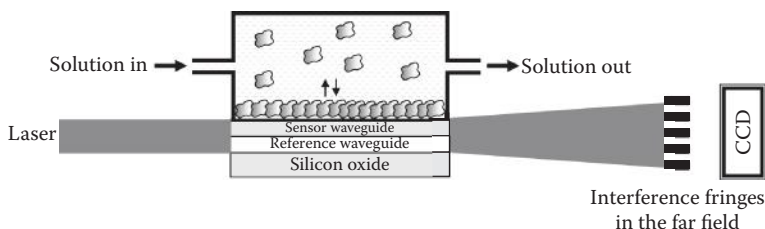
Interferometry is the study of the ways in which light waves interact or interfere with each other. Perhaps the most well-known example of early scientific interferometry is the diffraction pattern observed by Thomas Young in his double-slit experiments. Young noticed that if he shone a light beam on a screen that contained two closely spaced slits, then a characteristic interference pattern of alternating light and dark

bands could be observed on a second screen located behind the first, as shown in Figure 4.15. This interference pattern could be explained by appealing to the wave nature of light—the light bands represented regions where the light waves emerging from each slit interfered with each other constructively, whereas the dark bands were the result of destructive interference of the light waves.

It is possible to construct a waveguide-based interferometer analogous to Young's interferometer using a setup as shown in Figure 4.16. In this type of waveguide, there are two waveguide cores rather than one, and the cores are separated by a thin cladding region. The top waveguide core is termed the sensing waveguide and the bottom core is termed the reference waveguide. If a broad beam of light is shone on the edge of this type of “waveguide stack,” the light is totally internally



**Figure 4.15** The classic Young's double-slit experiments. Interference between the two wavefronts emerging from the double slit produce an interference pattern on a screen placed some distance away.



**Figure 4.16** Typical architecture of a DPI flow cell. Light enters the stacked waveguide and upon exiting generates an interference pattern in the far field due to a phase shift. The phase shift occurs due to the adsorption of material onto the sensor waveguide.

reflected through both of the core regions and emerges from the other side to generate an interference pattern, as shown in Figure 4.16.

This waveguide interferometer can be used to determine the refractive index of a small nanofilm if the top cladding region is replaced by a sensing region (generally an aqueous solution maintained by a fluidic cell, as shown in Figure 4.16). In this setup, light travels through the sensing waveguide at a speed that depends in part on the refractive index of the area within  $\sim 100$  nm of the sensing waveguide surface (i.e., within the evanescent field). Hence, any thin films or molecules that are adsorbed to the sensing waveguide surface change the speed at which the light travels through the sensing waveguide (because the adsorbed molecules change the refractive index of the sensing region within the evanescent field). On the other hand, light always travels through the reference waveguide at the same speed because the cladding regions on either side are fixed. And, since the relative speeds of the light emerging from the two waveguides affects the interference pattern produced, the interference pattern changes as a thin film is adsorbed to the sensor surface. More specifically, the relative positions of the light and dark bands (or fringes) shift as the film adsorbs to the surface. By using Maxwell's equations of electromagnetism and after applying some rather complex mathematics, the shifts in the fringe pattern can be used to calculate the refractive index of the film adsorbed to the surface of the sensing waveguide. Thus, a waveguide interferometer is able to characterize a thin nanofilm.

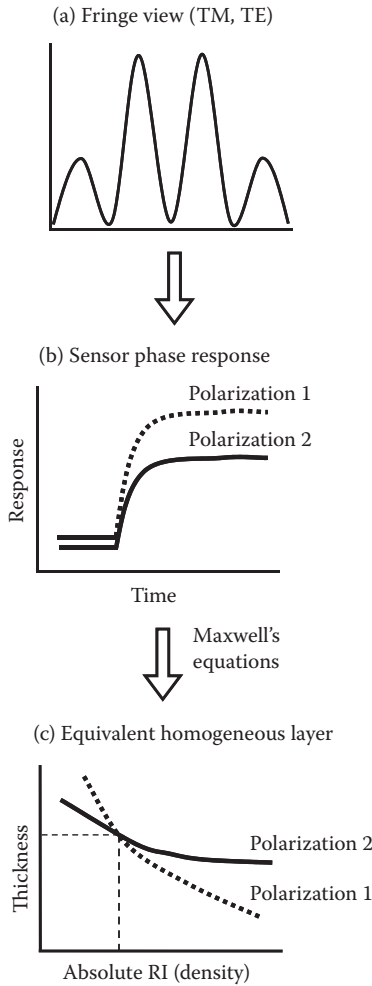
The refractive index measured by an interferometer such as described above is not the true refractive index of the film, but rather the *effective refractive index*, which is a complicated function of both the absolute refractive index and the thickness of the film. Incidentally, this kind of data is the same as that obtained using SPR, where the SPR angle change is a function of the change in absolute refractive index and the change in thickness of the film (see Equation 4.18 in Section 4.4.1). Normally the effective refractive index cannot be dissected into the absolute refractive index and thickness; however, if two different types of measurements of the effective refractive index are made, then it is possible to mathematically calculate both parameters. DPI makes these two measurements by using two different types of polarized light.

### 4.5.3 Principles of Dual Polarization Interferometry

DPI uses a waveguide interferometer setup as described above, only the light source is modified such that it alternately produces two different types of linearly polarized light via a polarizer switch (see Figure 4.16). One type of polarized light, named the transverse magnetic (TM) polarized mode, is composed of light waves with an electric field that oscillates perpendicular to the direction of the waveguide core, as shown in Figure 4.16. The other type, called the transverse electric (TE) mode, has light waves with an electric field that oscillates parallel to the direction of the waveguide. Incidentally, if we define the waveguide core as the plane of incidence, then the TM and TE modes directly correspond to the s- and p-polarized light modes described in our discussion of ellipsometry. Both of these polarization modes generate evanescent fields that extend into the cladding or sensing regions, but the fields produced by each mode are of different intensities and decay at different rates. Therefore, each polarized mode generates its own interference pattern on the detection screen and consequently each polarized mode provides a separate calculation of the effective refractive index.

It is important not to be misled by the direction of oscillation of the two polarized modes into thinking that no evanescent field would be generated by the TE mode, which oscillates parallel to the direction of the waveguide core. Regardless of polarization, the light waves still undergo total internal reflection (meaning the light beams still “bounce” off the cladding and/or sensing regions) and the two polarized modes still generate evanescent fields in the surrounding regions.

For the effective RI determined by each polarized mode, a large number of absolute refractive index and thickness values can be calculated that could possibly yield the observed effective RI, as shown in Figure 4.17. However, there is only one unique pair of absolute RI and thickness values that may generate the observed effective RI for both polarized modes. This pair represents the actual value of the absolute RI and the thickness of the film on the sensing waveguide surface. Therefore, by using two different polarized modes of light and by calculating the unique solution pair, DPI can be used to determine the actual RI and thickness of thin nanofilms. The technique is so sensitive that thickness changes of less than 1 angstrom are detectable. Furthermore, if one can



**Figure 4.17** (a) Typical representation of the fringe pattern observed in a DPI experiment. (b) The sensor phase response for the two polarization states. (c) Effective refractive index plots showing a unique solution for TE and TM modes. The point at which the curves cross gives the absolute refractive index and thickness of the film.

assume that the RI of the thin film is a linear function of the density of its contents (a good assumption for many thin films), then the RI can be manipulated to yield density  $\delta$  of the film according to

$$\delta = \frac{n_{\text{film}} - n_{\text{buffer}}}{dn_{\text{film}}/dc} \quad (4.19)$$

where  $dn_{\text{film}}/dc$  is the change in refractive index of the film per change in content density and  $n_{\text{film}}$  and  $n_{\text{buffer}}$  are the refractive indexes of the film and buffer (or solvent), respectively. In order to get the mass per unit area of the film, one need merely multiply the calculated density value by the average thickness of the film. Thus, DPI can be used to calculate the average density, mass, and thickness of the film simultaneously.

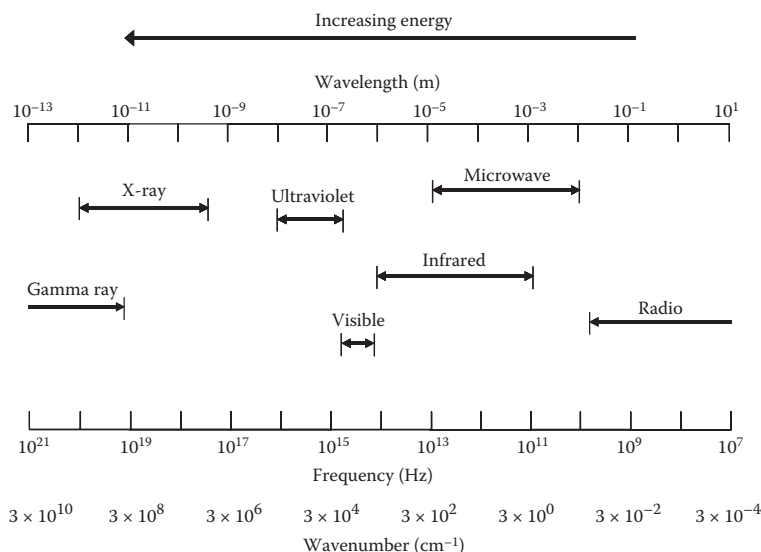
#### 4.5.4 Parameters of a DPI Instrument and Common Applications

A commonly used dual polarized interferometer is manufactured by Farfield Group Ltd, Manchester, UK. In their instrument, the DPI sensor surface is composed of silicon oxynitride, which allows for a wide range of covalent surface modifications through siloxane chemistry (discussed in Chapter 5). The flow cell that exists above the sensor surface has a volume of 2  $\mu\text{L}$ , and typical flow rates are on the order of tens of  $\mu\text{L}/\text{min}$ . Their instrument boasts thickness detection limits of  $<1$  angstrom and mass detection limits of  $100 \text{ fg}/\text{mm}^2$ . In addition, it can make multiple measurements per second, which allows it to monitor surface alterations in real time.

DPI has been widely marketed as a biosensor to measure the interactions between proteins and their substrates and to study the formation of membranes, supported lipid bilayers, and the interaction of surfaces with lipid vesicles. A non-biological application of DPI is the characterization of the physisorption of nanospheres to hard surfaces. Because it is able to simultaneously determine both thickness and RI (and consequently density and mass), DPI is also routinely used to monitor the changes in surface morphology of various thin films. For example, DPI can be used to estimate the shape of a protein adsorbed to a surface under conditions of differing pH.

### 4.6 SPECTROSCOPIC METHODS

Spectroscopy is the study of the ways in which electromagnetic radiation interacts with matter. Spectroscopy is an invaluable field of chemical analysis, and different spectroscopic methods are able to determine identity, concentration, and structural information of chemical compounds, among other useful information. Because the electromagnetic spectrum is so broad, different spectroscopic methods are employed for different regions of the spectrum. Figure 4.18 depicts a listing



Spectroscopy	Typical Range	Transition Studied
X-ray absorption and diffraction	0.1–100 Å	Inner electron energy states
UV-vis absorption and fluorescence	200–800 nm	Outer (bonding) electron energy states
IR absorption and raman scattering	0.8–300 μm (14000 to 30 cm <sup>-1</sup> )	Vibrational energy states
Nuclear magnetic resonance	0.6–10 m	Spin of nuclei in magnetic field

**Figure 4.18** The electromagnetic spectrum and a listing of common spectroscopic methods used to interrogate those regions of the spectrum. Note that the energy of light increases at smaller wavelengths and higher frequencies.

of the common regions of the electromagnetic spectrum as a function of wavelength and frequency. It also lists the most common spectroscopic methods employed in each region of the spectrum and the kinds of transitions those methods probe.

The following sections examine the application of a variety of spectroscopic techniques to the study of nanomaterials. In order to better understand these methods, it is appropriate to begin with a brief discussion of the interactions between light and matter.



#### 4.6.1 Interactions Between Light and Matter

As described earlier (see Section 4.3 on ellipsometry), in many instances light is best thought of as an oscillating electromagnetic wave with a characteristic energy  $E$ . The energy of a photon of this electromagnetic radiation is a function of its frequency  $\nu$  (or its wavelength  $\lambda$ ), and can be calculated by Einstein's famous equation

$$E = h\nu = \frac{hc}{\lambda} \quad (4.20)$$

where  $h$  is Planck's constant and  $c$  is the speed of light. This equation was introduced in Chapter 2; because of its importance we repeat it here.

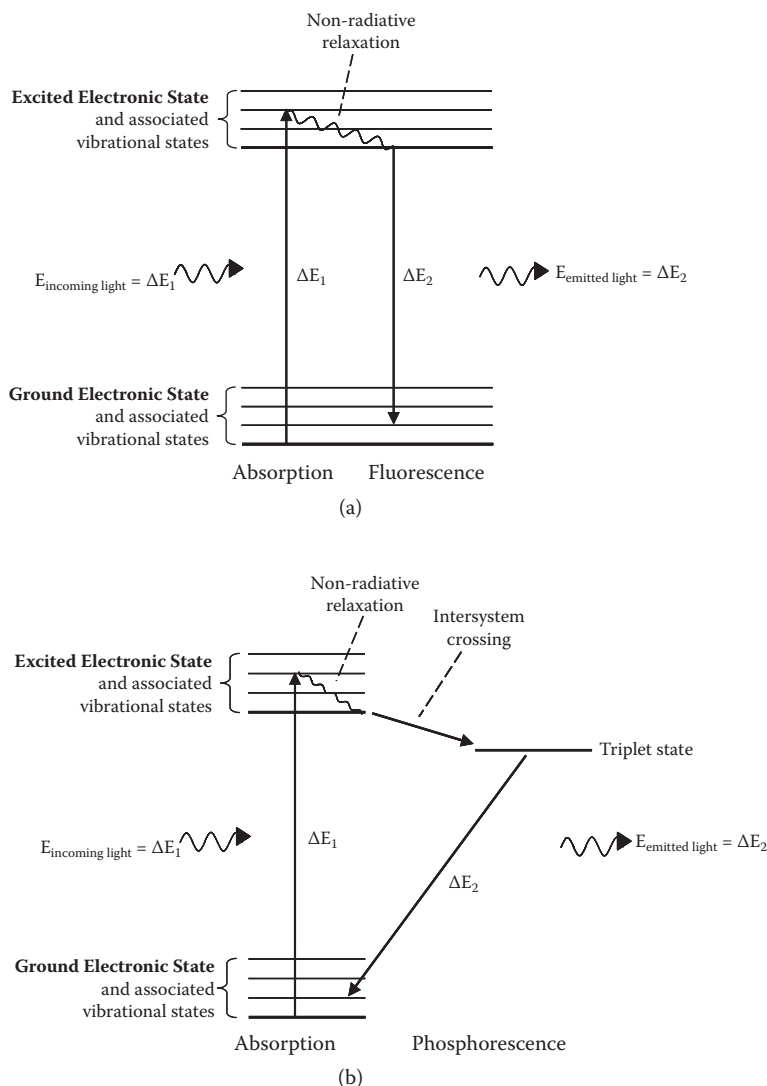
Molecules also have characteristic energy states, but it turns out that a given molecule cannot exist in any arbitrary energy state. Rather, the possible energy states for a molecule are quantized, meaning that the molecule is only allowed to exist in a finite number of discrete energy states. These discrete allowable energy states are the sum of several quantized aspects of the molecule, such as the energies of its electrons around their respective nuclei, the interatomic vibrations that exist in the molecule, and the rotations of the molecule around its center of mass. We can also say that a given molecule has quantized (or discrete) electronic, vibrational, and rotational energy states and that it can only exist at those energy states (or at a sum of those energy states). The lowest energy state of a molecule is termed the ground state and higher energy states are referred to as excited states. When energy is channeled into a molecule (say, by heating it or allowing it to absorb radiation), it may enter an excited state. Following this excitation, the molecule may relax from the excited state to a lower excited state or to the ground state. This relaxation is often accomplished by the emission of electromagnetic radiation (and consequently the frequency and the wavelength) of the emitted light is the exact difference between the upper and lower energy states. Using Einstein's equation (Equation 4.20), this process can be written as

$$E_1 - E_0 = h\nu = \frac{hc}{\lambda} \quad (4.21)$$

where  $E_1 - E_0$  is the difference in energy between the higher and lower energy states. Therefore, by measuring the wavelength

of the emitted light, one can calculate the energy difference between the higher and lower states of the molecule.

On the other hand, when electromagnetic radiation is shone upon a substance (rather than being emitted from it), several processes may take place (see Figure 4.19). Typically, the



**Figure 4.19** Energy diagrams depicting the transitions that occur during fluorescence (a) and phosphorescence (b). Note that the energies of the photons (either those being absorbed or those being emitted) are identical to the corresponding difference in energy states of the molecule.

radiation is transmitted through the matter at a velocity that depends upon the refractive index of the medium. However, if the frequency (or energy) of the incoming light happens to match the exact difference between two of the allowed energy states in a molecule, then the incoming light may be absorbed by the molecule. This outcome, in turn, promotes the molecule to an excited energy state and decreases the intensity of the light at that frequency according to the proportion absorbed by the molecule. This entire process is called absorption, and the absorption of different frequencies of electromagnetic radiation corresponds to different types of energy transitions in a molecule. For example, absorption of light in the ultraviolet and visible region of the electromagnetic spectrum typically causes transitions between electronic energy states of a molecule. Alternatively, absorbed infrared radiation causes a transition in the vibrational states of the molecule. It is important to emphasize that the incoming radiation is only absorbed by the molecule if the frequency of the radiation matches the exact energy difference between allowed energy states of the molecule. Otherwise, the radiation is only transmitted through the substance. Therefore, if one is able to measure the frequencies at which a molecule or substance absorbs the incoming radiation, then valuable information about that molecule can be determined, such as the type and strength of different bonds that exist in the molecule.

After the molecule has absorbed the incoming radiation, it ordinarily returns to the ground state after some short amount of time due to one of several relaxation processes. For some molecules, fluorescence is a common relaxation process in which the molecule relaxes by reemitting light, generally of a lower energy (longer wavelength) than the absorbed light. Fluorescence is typically observed in the ultraviolet-visible region of the electromagnetic spectrum and is depicted schematically in Figure 4.19. Nonradiative relaxation is the most common relaxation process, in which the excited molecule relaxes without reemitting any electromagnetic radiation. This process typically occurs in small steps by the conversion of the excited energy into kinetic energy through collisions with nearby molecules, producing heat. Nonradiative relaxation is also depicted in Figure 4.19.

Phosphorescence is another relaxation process that typically occurs in the UV-visible region. It is observed when an excited electron undergoes nonradiative intersystem crossing to a more rare and slightly more stable excited state called

a triplet state. When the excited electron then relaxes back to a state of lower energy, a photon is produced at a longer wavelength (smaller frequency) than the absorbed light. Phosphorescence is a much more rare phenomenon than fluorescence and is consequently less important to our discussion of spectroscopic techniques.

In summary, molecules may exist only at discrete, quantized energy states. By examining the ways in which a given molecule interacts with light, one can gather information about its energy states, thus providing valuable insight into the identity of the molecule, the strength and type of its chemical bonds, and the concentration of the molecule in the substance being studied. For our elementary purposes, those observations serve as the basis of spectroscopy. Let's consider several important types of spectroscopy and their applications to the study of nanomaterials.

## 4.6.2 UV-Visible Spectroscopy

### 4.6.2.1 Principles of UV-Visible Spectroscopy

UV-visible spectroscopy uses the transmission of visible and/or ultraviolet light through a sample to determine the presence and/or the amount of material that absorbs light within the sample. As mentioned in our discussion of the interactions between light and matter, absorption of a photon in the ultraviolet and visible wavelength range (~190 nm to 800 nm) typically produces an electronic transition within the absorbing molecule. The absorption of a photon increases the energy of the molecule, causing the promotion of the molecule out of a low-energy ground state electron configuration into an excited energy state. The region of a molecule that absorbs the light and undergoes the electronic transition is called the *chromophore*. For example, the amine group  $\text{-NH}_2$  absorbs light at a wavelength of about 190 nm. The absorption causes an electron from the lone pair on nitrogen to be excited into an antibonding molecular orbital. This transition is denoted as  $n \rightarrow \sigma^*$ . For our purposes, we consider only the amount of light absorbed and the corresponding absorption wavelength ( $\lambda_{\text{max}}$ ). The exact electronic nature of the transitions is less important. It should be noted that the actual absorption of chromophores in a given molecule depends on the exact electronic environment of the chromophore within the molecule. As a matter of convention, we also note that values for the UV-visible spectrum are typically reported in wavelengths of light (generally nm).

When absorption takes place in the sample, there is a decrease in the intensity, or radiant power, of the light beam passing through the sample. The transmittance  $T$  of the sample is defined as the ratio of the intensity of the beam after passing through the sample ( $I$ ) to the original intensity of the beam ( $I_o$ ). The absorbance  $A$  of the sample is defined as the negative logarithm of transmittance, as shown in Equation 4.22.

$$A = \log\left(\frac{I_o}{I}\right) = -\log T \quad (4.22)$$

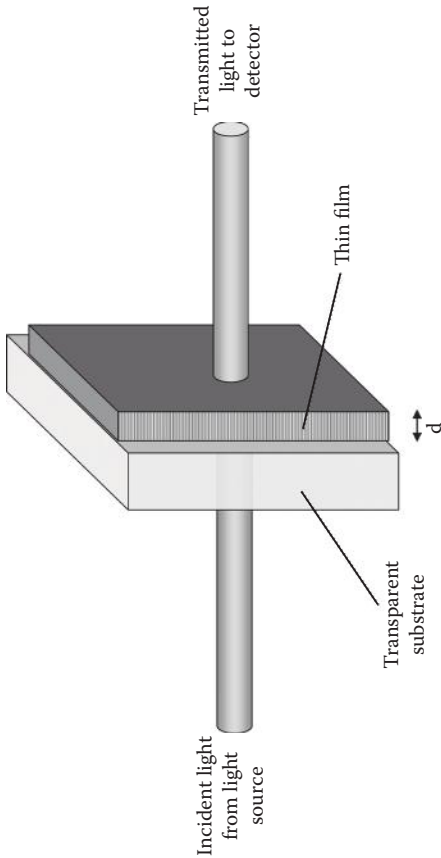
For samples with relatively low concentrations, the absorbance of the sample can be directly related to its concentration. This relationship is known as the Beer-Lambert law and is given in Equation 4.23.

$$A = c \epsilon l \quad (4.23)$$

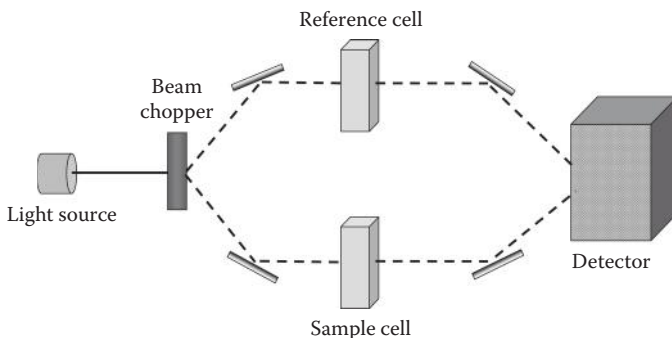
where  $c$  is the concentration of the sample,  $l$  is the sample path length, and  $\epsilon$  is a parameter known as the extinction coefficient and is also the molar absorptivity of the sample. The molar absorptivity is an important parameter that gives the probability of absorbing a photon for a particular molecule at a specific wavelength of light. Its value is specific to a particular chromophore and does not depend on concentration or path length. If the path length and  $\epsilon$  of a sample are known, then the density of the sample's contents can be calculated from the absorbance value. Alternatively, if the content density is known, then the path length can be calculated. In the context of a thin film (Figure 4.20), the path length is simply the film thickness ( $d$ ), and in this manner the thickness of a nanofilm can be determined using UV-vis spectroscopy.

#### **4.6.2.2 Setup of a UV-Visible Spectrophotometer**

At a basic level, a UV-vis spectrophotometer consists of a light source, a dispersive device, a sample holder, and a detector, along with several optical components. Common kinds of spectrophotometers are double-beam spectrophotometers, which use a beam splitter to allow the light beam to pass through a sample cell and a reference cell alternately throughout the measurement process. The beam is diverted into the reference cell by the splitter several times per second, and the



**Figure 4.20** Transmission mode UV-vis absorption spectroscopy can be used to determine the thickness of a thin nanofilm. Beer's law is used just as with bulk phase measurements, but the thickness of the film replaces the path length of the sample cell in the equation, as shown.



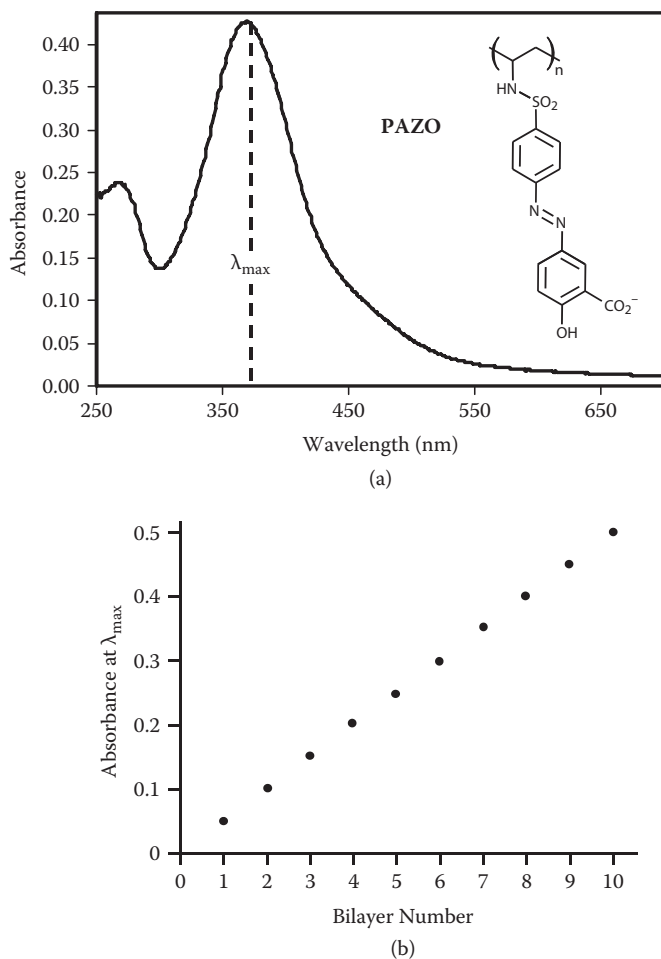
**Figure 4.21** A schematic of a common double-beam spectrophotometer used to make bulk phase measurements.

absorbance of the reference cell is measured and automatically compared with the absorbance of the sample cell to produce a more accurate value of the absorbance of the species being measured by correcting for fluctuations or variations in the lamp intensity over time. A schematic of a double-beam spectrophotometer is shown in Figure 4.21.

In practice, UV-visible spectrometers contain two different light sources, one to produce light in the UV range and one to produce visible light. The light sources are never run simultaneously. Instead, as the spectrophotometer measures the absorbance of a sample across a predetermined range of wavelengths, the light source switches from the visible source to the ultraviolet source. The light change is generally set to occur around 360 nm to allow for the highest light intensity at each wavelength.

#### 4.6.3 The Absorption of Visible Light by a Nanofilm

As an example of the application of UV-visible spectroscopy to characterize the growth of a nanofilm, consider the anionic polymer PAZO whose structure is shown in Figure 4.22. The PAZO monomer unit contains two phenyl groups bridged by an azo group ( $-\text{N}=\text{N}-$ ). This unit functions as a chromophore and absorbs light at about 360 nm due to a  $\pi \rightarrow \pi^*$  transition. Thus, an aqueous solution of PAZO appears orange because of the broad absorption of light centered at 360 nm. The polymer polyethylenimine (PEI) encountered previously contains amine groups that absorb light at about 190 nm. There is no absorption in the visible region and thus an aqueous solution of PEI is colorless. Both PAZO and PEI are commonly used to



**Figure 4.22** (a) The absorption spectrum of PAZO (structure shown in the inset).  $\lambda_{\max}$  represents the wavelength at which the absorbance is a maximum. (b) A plot of absorbance at  $\lambda_{\max}$  as a function of bilayer number. The bilayer represents a layer of polycation complexed with a layer of polyanion.

create multilayer films for a variety of applications. We will discuss these applications and the methods used to fabricate the multilayers in Chapter 5. About 1-nm thick layers of PEI and PAZO can be constructed on a glass substrate in a layer-by-layer fashion. These layers are held together by strong electrostatic interactions between the positively charged PEI and the negatively charged PAZO. Each PEI and PAZO pair in the film can be described as a “bilayer.” Figure 4.22(b) shows a

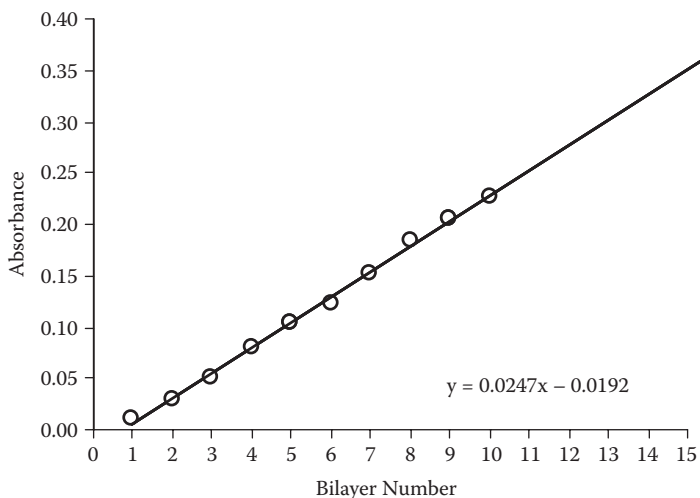


UV-visible spectrum of a PEI and PAZO multilayer film as a function of the bilayer number. Only the absorbance value at  $\sim 360$  nm is shown. In this range, only the chromophores associated with PAZO are able to absorb light (centered at a wavelength of 360 nm). The PEI is essentially transparent, and therefore the increase in the absorbance in Figure 4.22(b) reflects the increase in the amount of total PAZO in the film after each bilayer.

#### Example 4.6 Determining Absorbance Values

Estimate the slope of the line of the graph shown in Figure 4.22 and use it to predict an absorbance value of a film composed of 15 bilayers. How would you determine the concentration of PAZO in the film from the PAZO's molar absorptivity? What other technique would be useful in this determination?

**Solution** A linear fit to the data yields a slope of  $\sim 0.02$  (Figure 4.23). The intercept is close to zero. The slight negative number in the intercept is likely due to a baseline shift. Thus the equation of the line is  $y = 0.02x$ , where  $y = A$ . When  $x = 15$ , the absorbance ( $A$ ) is 0.75. From Beer's law,  $A = c\epsilon l$ , where  $l$  represents path length, or in this case the thickness of the film. We can estimate  $l$  using ellipsometry, and if we know  $\epsilon$  (molar absorptivity), we can obtain the concentration,  $c$ .



**Figure 4.23** The absorbance (at  $\lambda_{\max}$ ) data of a polyelectrolyte multilayer film. The bilayer represents a layer of polycation complexed with a layer of polyanion.

## 4.6.4 Molecular Fluorescence Spectroscopy

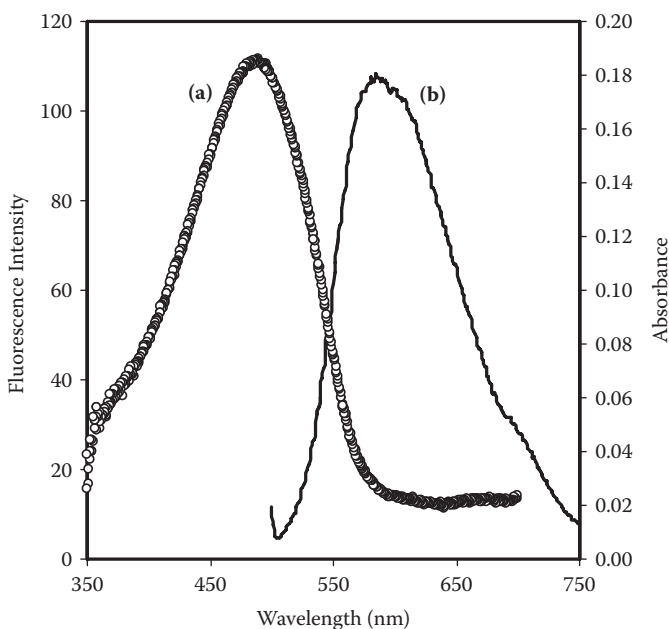
### 4.6.4.1 *Principles of Fluorescence and Fluorescence Quantum Yield*

As described in Section 4.6.1 on the interactions between light and matter, fluorescence is the process whereby a molecule that has been excited by the absorption of radiation relaxes to a lower energy state by emitting a photon of lower energy than the photon that was absorbed. Some types of molecules do not exhibit fluorescence; molecules that do fluoresce are called fluorophores. For most fluorophores, fluorescence is typically observed in the ultraviolet to visible range of the electromagnetic spectrum, meaning that it corresponds to transitions in the electronic state of the molecule (much like UV-vis absorption spectroscopy). As with UV-vis spectroscopy, the wavelength of light (generally nm) is used to describe fluorescence spectra rather than wavenumber ( $\text{cm}^{-1}$ ) or frequency (Hz) units.

Each fluorophore typically has a characteristic absorption profile, which is identical to the UV-vis absorption spectrum for the molecule. Many of the wavelengths at which the fluorophore absorbs UV-vis light excite the molecule to a sufficient degree such that it fluoresces upon relaxation. Wavelengths that produce fluorescence when absorbed are referred to as excitation wavelengths. However, not all excitation wavelengths cause the fluorophore to fluoresce to the same degree or even to produce fluorescent light of the same wavelength. Therefore, for each excitation wavelength, the fluorophore has a characteristic fluorescence emission profile, which is the range and intensity of wavelengths that are produced when that excitation wavelength is used to excite the fluorophore. Alternatively, for each emission wavelength there also exists an excitation profile, which is the range of wavelengths that produce fluorescence of that emission wavelength and the intensities of emission associated with each of those excitation wavelengths. Both emission and excitation profiles provide useful information about the fluorophore and its local environment, so fluorimeters can often be set to determine either the emission spectrum at a set excitation wavelength or the excitation spectrum at a set emission wavelength. It should be noted that because fluorescence-emitted photons are typically of lower energy (i.e., longer wavelength) than their excitation photon counterparts, it is not surprising that the emission spectrum for a given fluorophore usually occurs at longer wavelengths than its excitation spectrum. Typical

emission and excitation spectra for an optically active polymer are shown in Figure 4.24. The polymer is a derivative of poly(phenylene vinylene) (PPV), and such materials are being developed for applications in nano-photovoltaics.

As discussed previously (Section 4.6.1), there are several alternative relaxation processes that may occur in lieu of fluorescence, such as non-radiative relaxation and intersystem crossing. Furthermore, reabsorption of the fluorescent photon by a neighboring fluorophore may occur, especially at higher concentrations of fluorophore. Therefore, it is convenient to define a parameter that represents how often a given fluorophore successfully produces a fluorescent photon. Fluorescence quantum yield ( $\Phi_f$ ) is the usual measure of such a parameter and is defined as the ratio of photons absorbed to photons emitted for a given fluorescing species. As can be imagined, the correct determination of  $\Phi_f$  is especially important in the characterization of certain emissive nanomaterials such as those designed for photoluminescent devices.



**Figure 4.24** Typical absorption (a) and fluorescence emission (b) spectra of PPV, an optically active polymer. The fluorescence intensity is in arbitrary units. The molecules are excited at ~475 nm (the wavelength corresponding to the peak absorbance) and emits light at ~575 nm (the wavelength corresponding to the peak fluorescence).

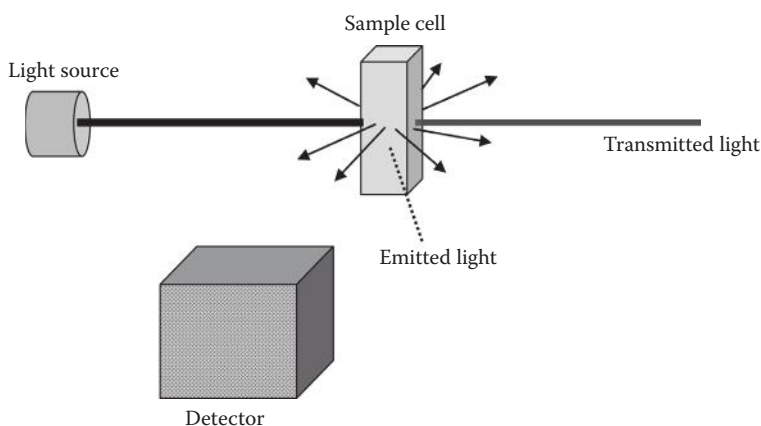
Determination of  $\Phi_f$  for a given material  $A$  can normally be accomplished in a rather straightforward manner by comparison with a standard material of known quantum yield  $\Phi_{std}$  according to

$$\Phi_{fa} = \Phi_{std} \frac{g_a n_a^2}{g_{std} n_{std}^2} \quad (4.24)$$

In this equation  $n_a$  and  $n_{std}$  are the refractive indexes of the solvents in which the material and the standard were dissolved, respectively.  $g_x$  is effectively defined as the derivative of the total fluorescence intensity of a fluorophore with respect to its absorbance of the excitation light (i.e., a measure of how the total fluorescence intensity changes with respect to absorbance). For low concentrations of the fluorophore, the relationship is usually linear (i.e.,  $g_x$  is constant), so  $g_x$  can be calculated by integrating the total fluorescence spectrum for a certain excitation wavelength and then plotting integrated fluorescence intensity versus absorbance of the excitation wavelength for a number of different concentrations.

#### 4.6.4.2 Setup of a Fluorometer for Bulk Phase and Thin Film Fluorescence Measurements

Figure 4.25 shows the typical setup of a fluorometer. The detector is usually placed at  $\sim 90^\circ$  from the incident light beam so as to not confuse light that has been transmitted through the sample with fluorescent radiation. Optical filters can



**Figure 4.25** A schematic of a typical fluorometer.

also be placed between the sample and the detector in order to ensure that only the fluorescence is detected. The fluorescence emission spectrum is gathered by shining a beam of a certain excitation wavelength on the sample and then scanning the output (emission) wavelength range with the detector while recording the intensity of fluorescence at each emission wavelength. Alternatively, the excitation spectrum can be obtained by setting the detector to monitor a specific emission wavelength and recording the intensity of fluorescence while scanning the input (excitation) wavelength range.

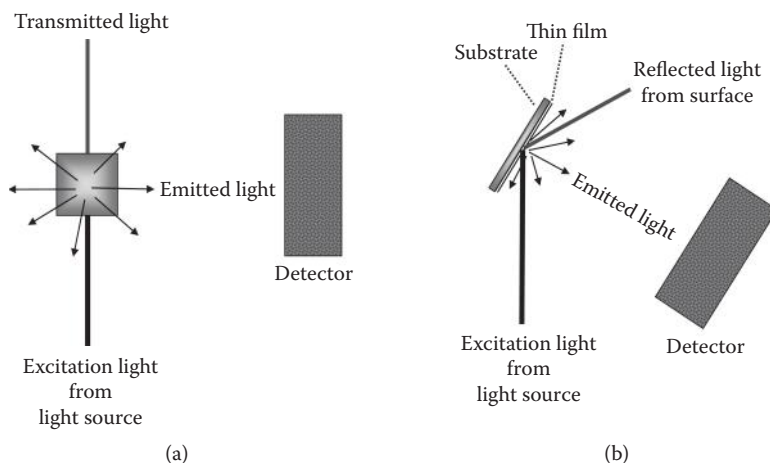
In some circumstances the intensity of the emission is too low to be recorded without amplification. In these situations it is best to use a photomultiplier tube. It is worth noting that photomultiplier tubes should be used only in cases of low-power radiation; otherwise, they will be damaged. A photomultiplier tube contains a photocathode surface that emits electrons when exposed to radiation. It also has multiple other electrodes, called dynodes. Each photoelectron that strikes a dynode causes many more electrons to be emitted. This creates an avalanche effect that results in millions of electrons for each initial photoelectron to be collected at an anode. The resulting current is measured in terms of voltage and is related to the wavelength being scanned.

For bulk phase measurements, the solution phase sample is generally placed into a fluorescence cuvette and fluorescence measurements are made at  $90^\circ$  to the incident beam. For thin films, a substrate on which the thin film has been built is placed so that the incident light shines upon the surface of the substrate at a large angle of incidence (i.e., more parallel to the surface) and fluorescence is detected at a different angle than the angle of reflection. Each of these two setups is depicted in Figure 4.26.

## 4.6.5 Vibrational Spectroscopy Methods

### 4.6.5.1 Introduction to Vibrational Modes

When a molecule absorbs infrared (IR) radiation, it is excited to a higher energy state, as described in Section 4.6.1. As with the absorption of other types of electromagnetic radiation, the IR absorption process is quantized and the molecule can absorb only particular frequencies of IR radiation. The absorption of IR radiation corresponds to a change in energy of approximately 20 kJ/mol. This is the amount of energy required to cause covalent bonds to stretch, bend, and twist, particular combinations of which are called the vibrational



**Figure 4.26** Methods for monitoring the fluorescence of samples in a cuvette in solution phase (a) or in a thin film (b). The detector is strategically placed so that it will intercept the maximum amount of fluorescent light without picking up any of the transmitted or reflected light.

modes of a molecule. Only those frequencies of IR radiation that match the natural vibrational frequencies of the covalent bonds of the molecule lead to IR absorption by the molecule. The energy absorbed leads to an increase in the amplitude of the vibrational motions of the bonds in the molecule. Furthermore, in order for a molecule to absorb IR radiation, it must possess a dipole moment that changes during the course of the vibration. A time-varying dipole moment means that the electron density between a covalent bond is changing. IR energy is transferred to the molecule when the frequency of this electrical oscillation is the same as the frequency of the incoming IR light. Therefore, symmetric molecules that have no dipole moment such as  $N_2$ ,  $H_2$ , and  $O_2$  do not absorb IR radiation. The natural frequency of vibration of a bond is given by

$$\bar{\nu} = \frac{1}{2\pi c} \sqrt{\frac{k}{\mu}} \quad (4.25)$$

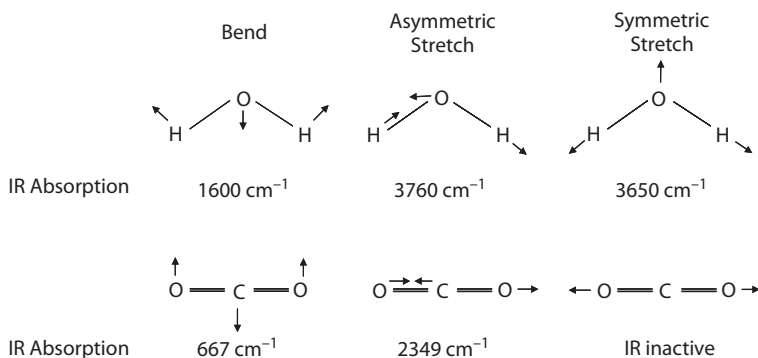
which is essentially the same as Hooke's law used to describe a spring undergoing harmonic oscillation. In this equation, the constant  $k$  is called the force constant of the bond and its units are N/m. The magnitude of  $k$  gives a direct measure of the stiffness of a covalent bond.  $\mu$  is known as the reduced

mass. For a simple diatomic molecule containing two atoms of mass  $m_1$  and  $m_2$ , the reduced mass is given by

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \quad (4.26)$$

As a simple example, consider the structure of  $\text{H}_2\text{O}$ . The central oxygen atom is  $\text{sp}^3$  hybridized, meaning that the molecule is bent with single bonds between the oxygen and hydrogen atoms. The molecule has a net dipole moment and is thus able to absorb IR radiation. Figure 4.27 shows three possible ways that  $\text{H}_2\text{O}$  can vibrate. When the two O-H bonds are simultaneously increasing and decreasing in length we say that the vibration is symmetric. If one O-H bond length increases while the other decreases, then we have an asymmetric vibration. The molecule can also undergo a bending motion. All three modes of vibration lead to changes in the dipole moment and therefore result in the absorption of IR radiation at characteristic frequencies. Since it is easier to bend the molecule than to stretch it, the bending mode has a relatively small  $k$  value. The implication is that, of the three observed IR absorption frequencies, the bending mode corresponds to the lowest frequency. If we replace the H atoms in  $\text{H}_2\text{O}$  with deuterium atoms ( $\text{D}_2\text{O}$ ), then the reduced mass of  $\text{D}_2\text{O}$  is greater than that of  $\text{H}_2\text{O}$ . According to Equation 4.25, all three IR absorption frequencies for  $\text{D}_2\text{O}$  are smaller than those for  $\text{H}_2\text{O}$ .

Let's consider the simple molecule  $\text{CO}_2$ . The central carbon atom is  $\text{sp}^2$  hybridized, meaning that  $\text{CO}_2$  is a linear molecule with double bonds between the carbon and oxygen atoms.



**Figure 4.27** The vibrational modes of  $\text{H}_2\text{O}$  and  $\text{CO}_2$  and the IR absorption associated with each vibrational mode. The symmetric stretch in  $\text{CO}_2$  does not produce a net change in the dipole moment of the molecule and is therefore IR inactive.

Since the molecule is linear and symmetric around the central C atom, CO<sub>2</sub> has a zero net dipole moment (the dipole moments from the two C=O bonds cancel each other out). When this molecule vibrates symmetrically, as shown in Figure 4.27, the net dipole moment remains zero. Thus, the symmetric vibration of CO<sub>2</sub> does not lead to the absorption of IR radiation. However, the asymmetric and bending modes of CO<sub>2</sub> do result in a net dipole moment, and therefore they lead to the absorption of IR radiation.

Organic molecules usually contain a variety of functional groups. Each functional group absorbs a specific IR frequency that is characteristic of that group. For example, the carbonyl group (C=O) absorbs around 1700 cm<sup>-1</sup>. The exact value depends on the local environment of the functional group, such as the identity of other atoms that are attached to the carbonyl functional group. Figure 4.28 shows a number of functional group vibrations and their characteristic IR absorption frequencies. As a matter of convention, we note that IR spectroscopists typically report IR values in wavenumbers (cm<sup>-1</sup>), a unit that is directly proportional to frequency  $\nu$  and defined as

$$\text{wavenumber} = \frac{\nu}{c} = \frac{1}{\lambda} \quad (4.27)$$

Functional Group	Group Frequency (cm <sup>-1</sup> )
-C-H (stretch)	2850-2960
=C-H (stretch)	3000-3100
≡C-H (stretch)	~3300
C=C (stretch)	1620-1680
C≡C (stretch)	2100-2260
-O-H (alcohols, H-bonded, stretch)	3200-3600
-O-H (carboxylic acids, H-bonded, stretch)	2500-3000
-N-H (stretch)	3300-3500
-N-H (bend)	~1600
C=O (stretch)	1670-1820
C≡N (stretch)	2220-2260
-S-H (stretch)	2550-2600
-S-S- (stretch)	470-620
Si-O-Si (stretch)	1020-1095
Si-O-C (stretch)	1080-1110
-N=N- (stretch)	1575-1630

**Figure 4.28** A listing of common functional groups and their IR frequencies.



where  $c$  is the speed of light (in  $\text{cm s}^{-1}$ ) and  $\lambda$  is the wavelength of light.

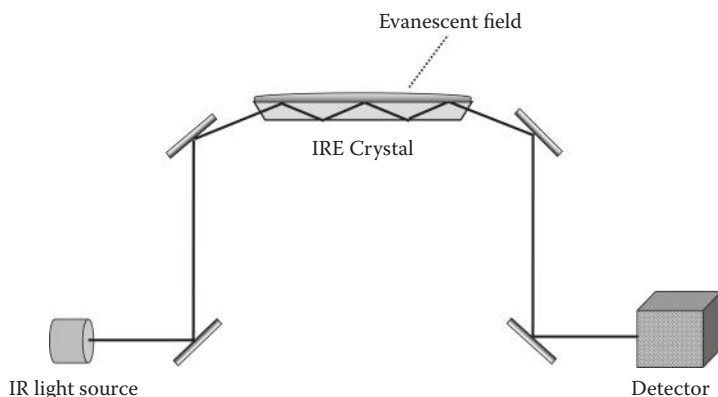
By plotting absorption or transmission versus incident IR frequency, one can map out the exact absorption frequencies for a sample containing the molecule of interest. This infrared spectrum provides a fingerprint of the molecular structure of the molecule.

Infrared spectroscopy is routinely used to elucidate the structure of organic molecules. We now consider a few specialized applications of IR spectroscopy suited for the analysis of nanomaterials.

#### **4.6.5.2 Attenuated Total Reflection IR Spectroscopy**

Attenuated Total Reflection Fourier Transform Infrared (ATR-FTIR) spectroscopy is a powerful spectroscopic technique used to investigate the structure of adsorbates confined to the solid-air or solid-liquid interface. The technique is essentially IR spectroscopy of molecules present at the surface of a solid. It offers several advantages over its counterpart transmittance-mode IR absorption spectroscopy. ATR-FTIR offers near surface selectivity with only a minimal amount of sample and allows for the detection of samples of mass on the order of nanograms. Furthermore, using an infrared polarizer, it is easy to determine the orientation of anisotropic (well-ordered) samples such as self-assembled surfactant or lipid monolayers. The sample can also be exposed to external conditions (e.g., various solvents or different pH conditions), making in situ studies feasible.

ATR-FTIR makes use of evanescent waves in a manner that is similar to DPI, as both are able to measure the intensity of the waves that are formed from total internal reflection within the surface-sensor medium. One significant difference is that ATR-FTIR monitors the absorption of IR light of an adsorbate rather than changes in its thickness and refractive index like DPI. A system of mirrors is set up so that IR light is shone upon a crystal called an Internal Reflection Element (IRE) in such a way that total internal reflection occurs, as shown in Figure 4.29. As with DPI, the total internal reflection of the IR beam results in the creation of an evanescent wave that penetrates the sample above the crystal and decays exponentially moving away from the crystal surface. (For a refresher on total internal reflection and evanescent waves, see Appendix A). For IR light, the penetration depth of the evanescent wave is typically on the order of a few micrometers (0.5–5  $\mu\text{m}$ ).



**Figure 4.29** A schematic diagram of an ATR setup for an FTIR spectrometer. The evanescent wave that is generated at the surface of the IRE can penetrate the overlying region up to several microns.

If the sample on the surface of the IRE happens to absorb the frequency of IR light that is being shone through the crystal, then the IR beam that emerges from the IRE and travels to the detector has a diminished intensity (or is *attenuated*). One can imagine that the adsorbate takes up the energy of the IR beam through the evanescent wave, but only at those IR frequencies that match the vibrational modes of the sample. Therefore, it is possible to obtain an IR absorbance spectrum for an adsorbate by scanning the entire IR spectrum and monitoring the frequencies at which the IR beam is attenuated upon its emergence from the IRE.

There are several common ways to set up an ATR-FTIR system to monitor the IR spectrum of a thin nanofilm. Perhaps the most straightforward method is to build the thin nanofilm on an external substrate and then clamp the substrate on top of the IRE, sandwiching the thin nanofilm between the substrate and the IRE. An IR absorbance spectrum of the nanofilm can then be obtained as described above. While this method is easy to implement, it does not allow for continuous monitoring of the thin nanofilm during its creation. To acquire that type of data, it is necessary to construct a flow cell on top of the IRE and build the thin nanofilm on the surface of the IRE itself. Using this setup, the IR spectrum can be monitored continuously, and changes that occur in the spectrum during the creation of the thin film can be recorded.

For the purposes of determining the IR spectrum of a thin nanofilm as described above, a horizontal ATR (HATR) setup is commonly employed. In this setup, the IRE is a parallel-sided

crystal plate with dimensions on the order of 1 cm by 5 cm. For an IRE of this size, the IR beam is typically reflected between 5 to 10 times at each surface, depending on the exact dimensions of the crystal and the angle of incidence. The upper face of the crystal is exposed to the thin film through a clamping method or a flow cell setup, as described. The most common crystal materials used for HATR setups are zinc selenide (ZnSe) and germanium. Diamond is a more robust material that can be used as an ATR crystal, although its higher cost is a drawback.

#### ***4.6.5.3 Reflection Absorption IR Spectroscopy***

Another surface-sensitive IR spectroscopy is Reflection Absorption Infrared Spectroscopy, or RAIRS. In RAIRS, infrared light is shone onto a metal surface at a grazing angle of incidence. The vibrational spectrum of molecules adsorbed onto the surface is obtained by comparing the intensity of the reflected light from a clean surface to one covered by a thin film. The vibrational modes that are observable in RAIRS are governed by a certain metal surface selection rule.

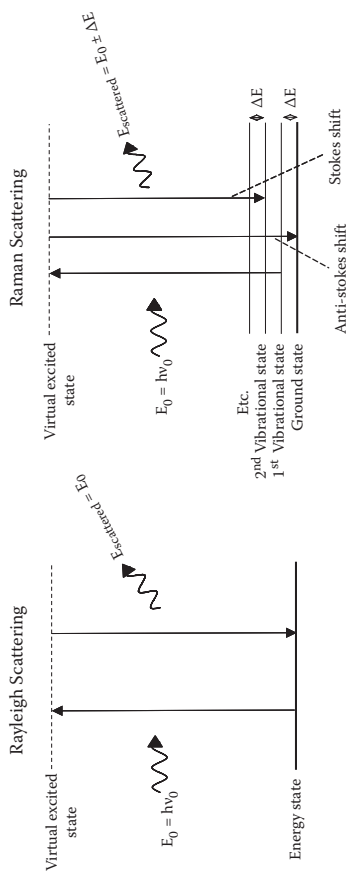
This states that only modes that have a component of their transition dipole moment, the complex vector quantity that includes the factors that change between two states, perpendicular to the metal surface appear in a RAIRS spectrum. Hence, the appearance of a RAIRS spectrum indicates that the transition dipole moments of specific functional groups in the film molecules have a component parallel to the surface normal.

### **4.6.6 Raman Spectroscopy**

#### ***4.6.6.1 Rayleigh and Raman Light Scattering***

In the opening discussion of the interactions between light and matter (see Section 4.6.1), we focused on processes such as absorption, emission, and fluorescence. However, another interaction called light scattering occurs commonly when light encounters nanoscale matter. In order to understand the scope and utility of Raman spectroscopy, a brief introduction to light scattering is provided.

When light passes through a solution or suspension, much of the light is transmitted directly through the solution, but some of the light is scattered in different directions. To better understand this scattering process, it is useful to model the scattering of light. Suppose that when a photon interacts with a molecule, the molecule is promoted to a virtual excited state as shown in Figure 4.30. This virtual state is short-lived, so we expect that the molecule almost immediately returns to its



**Figure 4.30** Energy diagrams modeling the Rayleigh and Raman scattering processes. A photon interacts with a molecule, promoting it to a non-quantized virtual state. The molecule quickly returns to a lower energy state, emitting a scattered photon. If the molecule returns to the same state in which it began, Rayleigh scattering has occurred. If the molecule returns to a higher or lower vibrational energy state, then Raman scattering has occurred.

ground state, reemitting the photon. However, the reemitted photon may not always be reemitted in the same direction as it was traveling before interacting with the molecule, which is scattering. It is important to note that this model of the scattering process should not be confused with absorbance and reemission of the photon. The incoming photon is not absorbed by the molecule according to our model. The promotion to the virtual excited state is a result of a momentary interaction with the photon and the entire process is not quantized (meaning it can happen for any energy of light). Absorbance only occurs when the frequency of the incoming light exactly matches the difference between allowed energy states of the molecule.

Most commonly, light is scattered elastically (or without losing or gaining any energy), in a process described as Rayleigh scattering. From Einstein's equation (Equation 4.20), we know that the energy of light is directly related to its frequency (or wavelength). Therefore, since Rayleigh scattering is elastic, the scattered light has the same frequency (or wavelength) as before it was scattered. However, due to the wave nature of light, Rayleigh scattering is a wavelength-dependent process, meaning that some wavelengths are scattered to a greater extent than others. Indeed, it is this aspect of Rayleigh scattering that accounts for the blue color of the sky. As light from the sun interacts with particles in the atmosphere, it undergoes Rayleigh scattering. Blue light is scattered more than the other wavelengths of light, because Rayleigh scattering is strongest for the shorter wavelengths of light that are closer in size to the air molecules responsible for scattering, so it appears as though the sky is a blue color. Rayleigh scattering is depicted schematically using our model of light scattering in Figure 4.30.

Raman scattering occurs when light is scattered inelastically by a molecule. In other words, Raman scattering happens when the scattered light is of a higher or lower energy after it has been scattered than before. This increase or decrease in energy is generally due to a change in the vibrational energy of the molecule. According to our model of light scattering, this type of Raman scattering might occur when the molecule relaxes from the virtual excited state to a vibrational state that is higher or lower in energy than the state at which the molecule was previously. In this case, the reemitted (or scattered) photon is a slightly different frequency (or wavelength) than before it interacted with the molecule. The shift in the frequency of the scattered light from its original value directly matches the energy difference between the vibrational states of the molecule.

In our diagram (see Figure 4.30), Raman scattering is represented as follows. Suppose the molecule happened to be in the first vibrational state above the ground state at the time of interaction with the photon. The molecule would be excited to the virtual excited state and then immediately relax, producing a scattered photon as described previously. Suppose, however, that the molecule would relax to either the second vibrational state or to the ground state rather than the first vibrational state. In this case, the reemitted photon would have either gained or lost energy according to the exact difference in energies between the vibrational states, represented by  $\Delta E$  in the figure. And, because the energy of light is a function of its frequency, the higher or lower energy light has a frequency that is slightly shifted from its original value and Raman scattering has occurred. Not only has the Raman scattered light changed direction, but it has also shifted its frequency.

Because the frequency shifts of Raman scattered light are a direct result of changes in vibrational states of the molecule, it seems reasonable to assume that one could gather similar information to IR spectroscopy by simply measuring the shifts in frequency of Raman scattered light. Indeed, this is the basic idea behind Raman spectroscopy, and frequency shifts in a Raman spectrum are directly analogous to IR absorption frequencies. In practice, shifts in frequency of Raman scattered light are monitored and those shifts are matched to specific bond vibrations within the molecule being studied. At a basic level, this allows one to identify the functional groups of the molecule being studied as well as their local chemical environment.

As a matter of naming convention, one should note that a frequency shift of Raman scattered light to a lower energy (or lower frequency) is called a Stokes shift. A frequency shift to a higher energy (or higher frequency) is called an anti-Stokes shift. Typically, Stokes shifts are more probable than anti-Stokes shifts, so their signals are stronger. As might be expected, Raman scattering is much less likely to occur than Rayleigh scattering for a given molecule and a given wavelength of light. Therefore, a Raman spectrum that maps intensity versus frequency of the scattered light has a centralized and strong peak that represents the Rayleigh scattering of the incident light. The Rayleigh peak is symmetrically surrounded by a distribution of much smaller peaks that represent the Stokes lines at the lower frequencies and the anti-Stokes lines at the higher frequencies. This peak is often blocked from the detector by a filter.

#### **4.6.6.2 Surface Enhanced Raman Spectroscopy**

Typically, the effects of Raman scattering are very weak. Under normal conditions, the intensities of the Raman scattered peaks in a spectrum are  $\sim 0.001\%$  of the incident light. Therefore, if a Raman spectrum is desired for an analyte with a relatively small concentration (such as for a thin film deposited on a surface), some method of signal enhancement must be employed in order to gather any useful information.

One method that is commonly used to enhance the Raman signal is adsorption of the analyte onto a rough metal surface or colloidal particle to obtain the Raman spectrum from the adsorbed species. Most often silver is used, but gold and copper have also been observed to produce similar effects. This type of signal enhancement is called surface enhanced Raman spectroscopy (SERS). SERS is a convenient method that allows amplification of the Raman signal while simultaneously probing the region near a metallic surface, thus allowing Raman spectra to be obtained for low quantities of a substance that are adsorbed to a metallic surface.

The exact mechanism by which adsorption to a metallic surface produces enhancement of the Raman signal is not fully understood. Two predominant theories have emerged to explain the effect. They are called the electromagnetic enhancement theory and the chemical enhancement theory, respectively. Each theory is able to explain some observations, but not others.

The electromagnetic theory claims that the generation of surface plasmons by the incident light is the means by which Raman scattering is enhanced. As may be recalled from our discussion of SPR (Section 4.4), a surface plasmon is a collective oscillation of the electrons at the surface of a conducting metal (typically gold or silver) that is induced when light is shone upon its surface. The generation of surface plasmons greatly enhances the electromagnetic field near the metallic surface and therefore would be expected to enhance the strength of the Raman scattered light. This, then, is the fundamental idea behind the electromagnetic enhancement theory. The incident light produces surface plasmons, which enhance the electromagnetic field near the surface, which in turn produces a stronger Raman signal.

The chemical enhancement theory appeals to a charge-transfer mechanism to explain the enhancement of Raman signal upon adsorption to a metallic surface. The charge-transfer

mechanism essentially states that new electronic states are made available to the molecule due to its interaction with the metallic surface. These new allowable electronic states can be of intermediate energy between the highest occupied molecular orbital of the molecule and its lowest unoccupied molecular orbital. Hence, the new electronic states can serve as resonant intermediaries in the process of Raman scattering. Therefore, excitations of the molecule that use the charge-transfer mechanism occur at much lower energies than are normally required for the molecule to be excited. Thus, Raman scattering can happen more readily and the signal is enhanced.

SERS is typically carried out on a “rough” metallic surface, with roughnesses on the order of 10–100 nm. Such surfaces are commonly manufactured by sputtering or evaporation of the metal onto a substrate or by roughening of a metallic electrode surface during oxidation-reduction cycles. Colloidal particles of the metal are also used to produce SERS. These colloids may either be suspended in solution or adsorbed onto a substrate. On these types of “rough” metallic surfaces, the SERS effect extends up to tens of nanometers from the surface, allowing for the effective probing of ultra-thin nanofilms.

## 4.7 NONLINEAR SPECTROSCOPIC METHODS

### 4.7.1 An Introduction to Nonlinear Optics

Nonlinear optics deals with the interaction of intense light with matter. The light source is usually a pulsed laser that can generate radiation with a strong electric field strength. Although the theory of nonlinear optics is beyond the scope of this book, a cursory treatment is provided to facilitate understanding of how nonlinear optical effects can be used to probe the properties of molecules in nanoassemblies.

This chapter has focused on linear optical effects such as reflection, refraction, absorption, and interference. These effects are observed from all light sources, regardless of intensity. Linear optical effects are based on a linear relationship between an oscillating electric field and an induced dipole moment in a molecule. Thus, when the oscillating electric field of light interacts with a molecule, the electron cloud in the molecule also begins to oscillate. This electron density oscillation sets up an oscillating dipole moment in the molecule. The strength of this dipole ( $\mu$ ) depends linearly on



the strength of the incoming electric field ( $E$ ) according to Equation 4.28:

$$\mu = \alpha E \quad (4.28)$$

The constant  $\alpha$  is known as the molecular polarizability. It measures the degree of electron density distortion due to the electric field. One consequence of this equation is that when light of frequency  $\omega$  interacts with a material, the frequency is unchanged. For example, when incident light of frequency  $\omega$  bounces off a surface, the reflected light will also be of frequency  $\omega$ . Now Equation 4.28 is actually a power series equation, with quadratic and higher order terms:

$$\mu = \alpha E + \beta E^2 + \gamma E^3 + \dots \quad (4.29)$$

The higher order terms become important only if  $E$  is large. The constants  $\beta$  and  $\gamma$  are known as the first and second hyperpolarizability, respectively. These constants are essentially zero when ordinary low-intensity light interacts with a molecule. However, when light from an intense pulsed laser is used, these nonlinear terms become significant. Equation 4.29 is the general equation describing the effect of light on a molecule. It is important to realize that all materials are nonlinear; it is a matter of the magnitude of the perturbing electric field that is needed to set in the anharmonicity and to drive it to nonlinear behavior. Extending this treatment to a bulk material, a similar equation is obtained, where the dipole ( $\mu$ ) is now replaced by the average polarization ( $P$ ) of the bulk material:

$$P = \chi_1 E + \chi_2 E^2 + \chi_3 E^3 + \dots \quad (4.30)$$

The constants  $\chi_1$ ,  $\chi_2$ , and  $\chi_3$  are known as first-, second-, and third-order susceptibilities. They are related to the sum components of the corresponding hyperpolarizabilities averaged over orientations of the molecules in the bulk material. The complete mathematical descriptions of the various  $\chi$  terms are ignored here. We can separate each polarization term in Equation 4.30, such that  $P_1 = \chi_1 E$  represents the linear polarization, and  $P_2 = \chi_2 E^2$  represents the second-order nonlinear polarization, etc. In fact, many of the nonlinear optical effects used to study nanomaterials are based on the second-order nonlinear polarization. We focus on this term.

Let's consider two incident intense laser sources with frequencies  $\omega_1$  and  $\omega_2$  incident on a material. These two frequencies have two oscillatory electric fields (with fields  $E_1$  and  $E_2$ ) simultaneously acting on the material, given by the equation

$$E = E_1 \cos \omega_1 t + E_2 \cos \omega_2 t \quad (4.31)$$

Substituting this equation into the expression for the second-order nonlinear polarization,  $P_2 = \chi_2 E^2$ , yields the following interesting result:

$$P_2 = \frac{\chi_2}{2} \left\{ (E_1^2 + E_2^2) + (E_1^2 \cos 2\omega_1 t) + (E_2^2 \cos 2\omega_2 t) + 2E_1 E_2 \cos(\omega_1 + \omega_2)t + 2E_1 E_2 \cos(\omega_1 - \omega_2)t \right\} \quad (4.32)$$

This is an interesting result because of the terms  $2\omega_1$ ,  $2\omega_2$ ,  $\omega_1 + \omega_2$ , and  $\omega_1 - \omega_2$ , which tell us that when two intense light frequencies ( $\omega_1$  and  $\omega_2$ ) interact with a material, we can produce resulting light of doubled frequencies ( $2\omega_1$ ,  $2\omega_2$ ), and even resulting light of combined ( $\omega_1 + \omega_2$ ) and subtracted ( $\omega_1 - \omega_2$ ) frequencies. The doubled frequency production is known as second-harmonic generation (SHG), and the latter two products are usually referred to as sum-frequency and difference-frequency generation (SFG and DFG). Two photons can pass through a material and combine in energy to produce a single photon. If we had only one frequency of light, then two photons would combine to produce a single photon of twice the energy (and hence frequency) of the incident photon. In this case, there is no distinction between SHG and SFG. We are combining two photons to produce one, so the conversion efficiency of this process has a theoretical upper limit of 50%.

Not all materials generate a second-order polarization, and those that produce this effect do so at varying efficiency. The key determining factor is  $\chi_2$ , the second-order susceptibility. The mathematical properties of  $\chi_2$  depend on factors such as molecular orientation, and these properties ultimately determine the nonlinear optical conversion efficiency. The properties of  $\chi_2$  are discussed in the next section.

#### Example 4.7 Combining Photon Energies

A pulsed laser light source composed of two wavelengths, (a) green light (500 nm) and (b) infrared light (900 nm), is passed through a

nonlinear optically active material. Determine the color of the second harmonic and sum-frequency light emitted from the material.

*Solution* Convert all wavelengths to frequencies:

$$(a) \quad \nu = \frac{c}{\lambda} = \frac{2.998 \times 10^8 \text{ ms}^{-1}}{500 \times 10^{-9} \text{ m}} = 5.996 \times 10^{14} \text{ Hz}$$

$$(b) \quad \nu = \frac{c}{\lambda} = \frac{2.998 \times 10^8 \text{ ms}^{-1}}{900 \times 10^{-9} \text{ m}} = 3.331 \times 10^{14} \text{ Hz}$$

Combine the frequencies:

SHG:

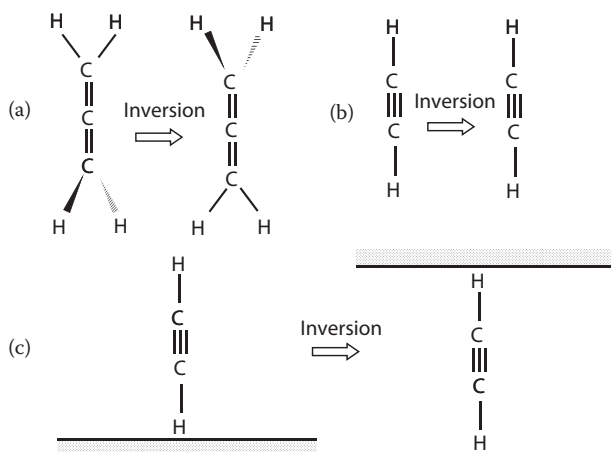
$$(a) 2(5.996 \times 10^{14}) = 1.199 \times 10^{15} \text{ Hz, or } \sim 250 \text{ nm (deep UV light)}$$

$$(b) 2(3.331 \times 10^{14}) = 6.662 \times 10^{14} \text{ Hz, or } \sim 450 \text{ nm (blue light)}$$

$$\text{SFG: } 5.996 \times 10^{14} + 3.331 \times 10^{14} = 9.327 \times 10^{14} \text{ Hz, or } \sim 320 \text{ nm (UV light)}$$

Two factors determine the magnitude of the second-order susceptibility. The first is symmetry, and the second is the permanent dipole moment in the molecule (force required to polarize the molecule). Generally speaking, conjugated molecules are easier to polarize than non-conjugated molecules because the electrons are delocalized along the carbon backbone. Furthermore, if a conjugated molecule possesses electron donor and acceptor groups on opposite ends, it makes it even easier to polarize the molecule. Such molecular configurations increase the nonlinear second-order effect.

$\chi_2$  is zero in a centrosymmetric environment. If a molecule possesses inversion symmetry,  $\chi_2$  is reduced to zero. Inversion involves passing each atom in a molecule through the center of the molecule and placing it on the opposite side of the molecule. Let's consider the ethylene and ethyne. After inversion, the ethylene molecule looks different, but the position of the hydrogens is the same for ethyne after inversion. We say the ethylene is non-centrosymmetric, or lacks inversion symmetry. Interestingly, a molecule at an interface always lacks inversion symmetry, even ethyne (Figure 4.31). Therefore, surfaces and interfaces are inherently non-centrosymmetric regions, and molecules confined to such regions produce SHG and SFG light. This fact allows SFG and SHG to be used as surface-specific probes to investigate molecular structure

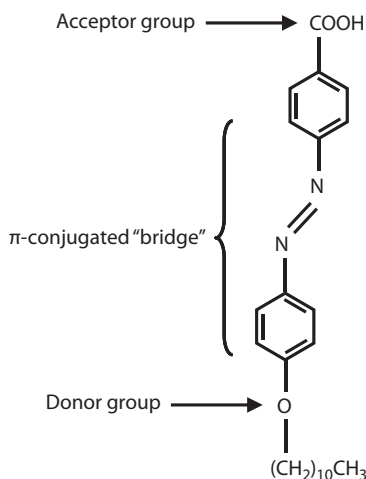


**Figure 4.31** An inversion operation on simple molecules. (a) Ethylene does not have inversion symmetry because the positions of all the hydrogen atoms are changed after the operation. (b) All atoms in ethyne are unchanged after an inversion. This molecule has inversion symmetry. If ethyne is on a surface, as in (c), the center of symmetry is broken. Surfaces represent non-centrosymmetric regions where there is no inversion symmetry.

at interfaces. Many nanomaterials are supported on surfaces and can be probed using such nonlinear optical techniques.

#### 4.7.2 Second-Harmonic Generation

SHG involves the conversion of frequency  $\omega$  to  $2\omega$ . The effect is strongest in non-centrosymmetric regions and in molecules that are highly polarizable. Consider an amphiphile containing a highly polarizable headgroup (Figure 4.32). The polarizable headgroup can be considered as a nonlinear optically active (NLO) chromophore. The NLO-active chromophore in such molecular building blocks contains an electron acceptor (A) and an electron donor (D), bridged by a conjugated azobenzene unit. This D- $\pi$ -A system is crucial for a strong NLO response. In Figure 4.32, the acceptor group is a carboxylic acid group and the electron donor is the oxygen atom. A monolayer comprised of this molecule produces a highly oriented assembly with a large net dipole. These conditions meet the requirement of a strong SHG signal. An additional layer can be added to the assembly either with the same orientation (asymmetric assembly) or with the opposite orientation (symmetric assembly). The latter reduces the SHG signal because of the centrosymmetric nature of the symmetric bilayer.



**Figure 4.32** An amphiphile containing an NLO-active headgroup. This moiety is comprised of an electron donor (the oxygen atom) connected to an electron acceptor (the carboxylic acid group). The azobenzene group serves as a  $\pi$ -conjugated bridge between the electron donor and acceptor groups. This headgroup is highly polarizable.

Constructing nanoassemblies that have a large NLO response is important to a number of technologies, including optoelectronics (using photons instead of electrons in electronic devices), communications, and data storage. Materials with efficient frequency doubling properties can also be used in lasers to generate light of different wavelengths. Self-assembly methods can be used to guide molecular building blocks into preferred orientations that yield new materials with NLO-active properties. Even though this field is leading the development of next-generation electronic and computational breakthroughs, fundamental limitations exist in fabricating functional nanoassemblies for NLO applications. These limitations arise from the inherent difficulty in obtaining asymmetrically oriented NLO chromophores in layered thin-film assemblies. As a result, the NLO response does not increase quadratically with the film thickness, rendering the assembly useless. Layer-by-layer methods, covered in Chapter 5, seem to be the most promising approaches to creating multilayered films containing NLO blocks throughout the 3-D assembly. An ideal NLO-active film would contain hundreds of discrete layers of highly oriented molecules self-organized on a solid support. In the case of NLO materials, if electron donor-acceptor chromophores were organized

in such a way that the dipole moments were oriented in the same direction, then a quadratic increase of the optical second-harmonic signal would be observed with increasing film thickness (number of bilayers) as predicted by Equation 4.33, where  $L$  is the film thickness,  $I_\omega$  is the intensity of incident light,  $n_\omega$  is the film's refractive index at the incident light frequency, and  $n_{2\omega}$  is the refractive index at the doubled frequency.

$$I_{2\omega} = \frac{2\omega^2 d_{eff}^2 L^2}{c^3 \epsilon_o n_\omega^2 n_{2\omega}} (I_\omega)^2 \frac{\sin^2(\Delta k L / 2)}{(\Delta k L / 2)^2} \quad (4.33)$$

The Langmuir-Blodgett (LB) method is a commonly used technique to organize amphiphilic molecules into densely packed 2-D structures at the air-water interface. The 2-D structure can then be conveniently transferred to a substrate. This method is discussed further in Chapter 5. By repeating the transfer process, ordered 3-D nanostructures of desired thickness can be obtained. If asymmetric structural order is retained throughout the assembly, one can then achieve a sum effect of the functionality represented by an individual layer. However, the idea of relying on the sum effect of individual LB layers to obtain an enhanced NLO functionality in 3-D materials is seriously challenged by studies that demonstrate a certain level of intermixing between layers. Intermixing randomizes the chromophore orientation, resulting in a reduction of the SHG intensity. We are interested in an asymmetric structure with a large optical nonlinearity. This type of deposition is usually accomplished using two different species (discussed further in Chapter 5). Disorder in LB assembled materials is due to the presence of kinked alkyl tails within the assembly, resulting in a decreasing packing density. Furthermore, disorder may also arise from a lack of in-plane ordering due to unfavorable dipole-dipole interaction. In a typical LB assembly involving two different species, both materials from a given deposition cycle may occupy defects in underlying layers. Materials of either species have a tendency to spread in a direction perpendicular to the substrate, leading to enhanced interpenetration, despite the maintenance of equivalent bilayer thicknesses over many layers. In other words, equivalence in deposition amount and thickness over many layers may indicate similar net surface conditions in each layer, but care must be taken in using the notion of a

similar surface to infer consistency in molecular order within layers. Even disordered systems can display seemingly regular deposition patterns—a film with disorder on the molecular level may appear markedly ordered at the macroscale. Such considerations are far reaching in a field largely focused on not only the order, but also the chromophore orientation at the molecular level.

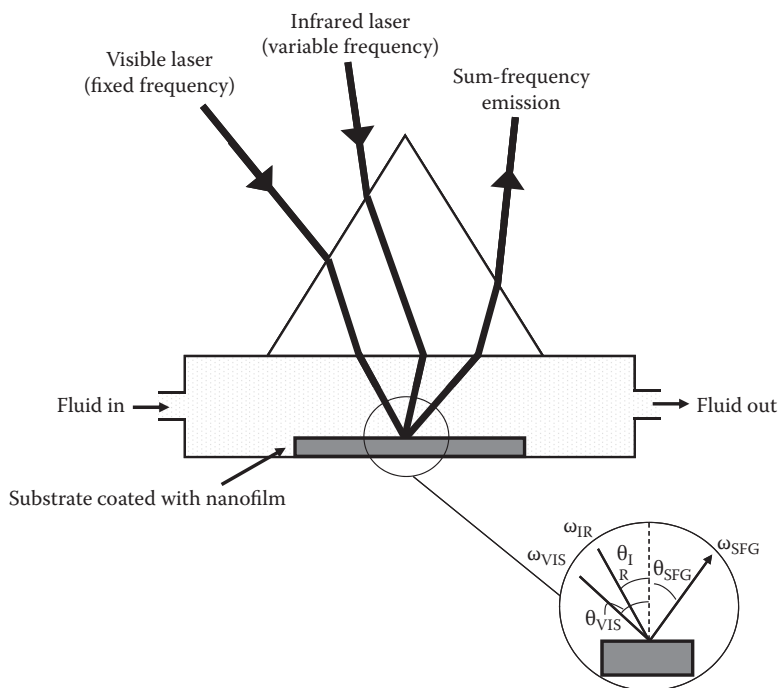
In addition to measuring the NLO response from a nanomaterial, SHG can also be used to follow the assembly process in real time, obtain the density of NLO-active chromophores within the assembly, and even image ordered domains in nanofilms. By obtaining the relative amount of SHG generated from s- and p-polarized light, the orientation of the NLO-active chromophore with respect to the surface normal can also be calculated.

#### 4.7.3 Sum-Frequency Generation Spectroscopy

Photons of two different frequencies ( $\omega_1$  and  $\omega_2$ ) can combine to generate light at the sum frequency ( $\omega_{\text{SFG}} = \omega_1 + \omega_2$ ). Similar to SHG, SFG is produced from non-centrosymmetric environments. SFG is routinely used as an IR spectroscopic method to obtain molecular structure information of nanofilms confined at interfaces. In this technique one of the two frequencies is fixed at a visible wavelength, typically 532 nm ( $\omega_{\text{vis}}$ ). The other frequency is variable in the IR region ( $\omega_{\text{IR}}$ ). The two light beams are overlapped on a surface and the intensity of the SFG beam is measured. It turns out that this intensity is greatly enhanced when  $\omega_{\text{IR}}$  is in resonance with a vibrational mode of the molecule at the surface. Thus, a plot of the SHG intensity versus  $\omega_{\text{IR}}$  provides a vibrational spectrum of molecules on the surface. It is worth noting that in contrast to ATR-FTIR spectroscopy, SFG spectroscopy is surface specific, whereas ATR methods probe regions within as much as a few microns from the surface. A theoretical expression for the magnitude of the SFG signal is given by Equation 4.34:

$$I_{\text{SFG}} = 128\pi^3 \left( \frac{\omega_{\text{SFG}}}{hc^3} \right) |K_{\text{SFG}} K_{\text{vis}} K_{\text{IR}}|^2 |\chi_2|^2 \left( \frac{I_{\text{vis}} I_{\text{IR}}}{AT} \right) \quad (4.34)$$

The signal intensity,  $I_{\text{SFG}}$ , is the number of SFG photons produced per laser pulse and depends on the beam intensities ( $I_{\text{vis}}$  and  $I_{\text{IR}}$ ), the laser pulse length ( $T$ ), the area of overlap of the



**Figure 4.33** Typical geometry of an SFG experiment in which a prism is used to couple the various laser beams onto a surface. The various angles of the IR, visible, and SFG beams are indicated in the circle.

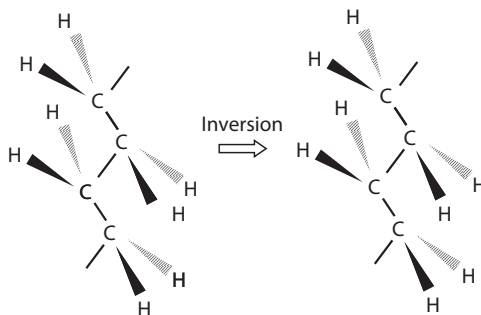
beams ( $A$ ), the various geometric Fresnel factors ( $K$ ), and the second-order susceptibility.

In practice, in order to generate SFG light, precise incidence angles of the fixed frequency visible and tunable IR light have to be used. Furthermore, the emitted SFG light is observed at a precise angle. This is known as phase matching, and the appropriate angles can be calculated using Equation 4.35. Figure 4.33 shows the typical geometry of an SFG experiment, indicating the various angles shown in Equation 4.35.

$$\omega_{\text{SFG}} \sin \theta_{\text{SFG}} = \omega_{\text{vis}} \sin \theta_{\text{vis}} - \omega_{\text{IR}} \sin \theta_{\text{IR}} \quad (4.35)$$

SFG spectroscopy has been used extensively to study the conformation of alkyl chains in monolayers assembled at interfaces. If the alkyl chain is in an all-trans conformation, then all the  $\text{CH}_2$  groups are in a locally centrosymmetric environment, as illustrated in Figure 4.34. The presence of a gauche defect breaks this centrosymmetry and renders the

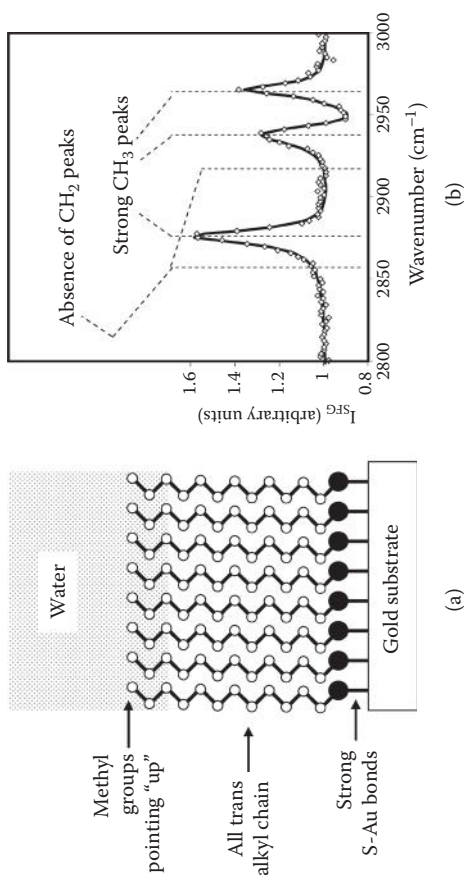




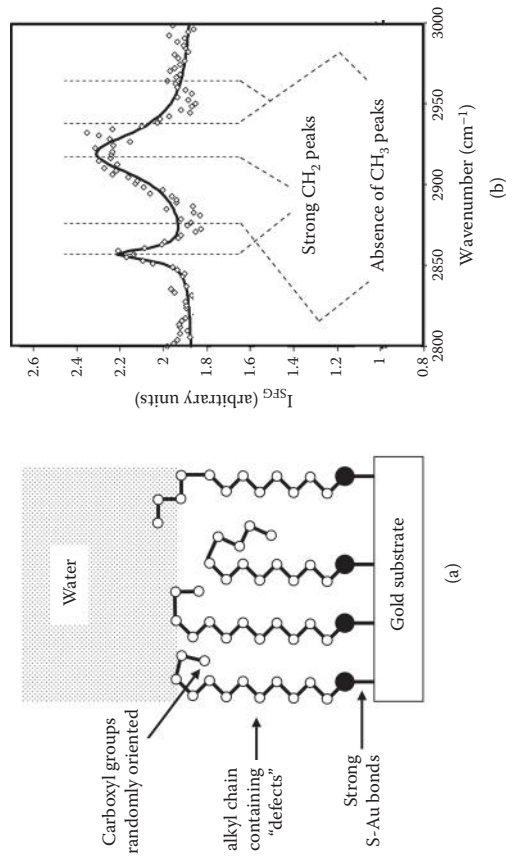
**Figure 4.34** Portion of an all-trans alkyl chain. The  $\text{CH}_2$  are locally centrosymmetric because after an inversion operation, the positions of all atoms remain unchanged. Therefore, these groups in this conformation will not produce SHG or SFG light. A “kink” in the chain or a gauche defect will create a non-centrosymmetric environment of  $\text{CH}_2$  groups, rendering these groups SFG active.

$\text{CH}_2$  sum-frequency active. The power of SFG spectroscopy is illustrated by the following example. Consider a monolayer of dodecane thiol (DDT) chemisorbed to a gold surface (Figure 4.35) and under water. Details of this type of monolayer are discussed further in Chapter 5. This monolayer is tightly packed with all the  $\text{CH}_2$  groups in an all-trans conformation. Only the terminal  $\text{CH}_3$  groups are non-centrosymmetric. The SFG spectrum of DDT shows only features that are attributed to the vibrational modes of the terminal methyl group. In comparison, Figure 4.36 shows the SFG spectrum of a mercaptododecanoic acid (MDA). This molecule is structurally similar to DDT, except that it contains a large carboxylic headgroup instead of a  $\text{CH}_3$  group. The SFG spectrum of MDA shows no  $\text{CH}_3$  features, but does show prominent vibrational features that arise from  $\text{CH}_2$  groups. The presence of the methylene features indicates a significant amount of gauche defects in the alkyl chains of the MDA monolayer, which is not surprising since, compared to DDT, MDA has a bulky headgroup that prevents the monolayer from packing tightly. This characteristic leads to a greater amount of space for the alkyl chain to occupy, resulting in the formation of kinks in the hydrocarbon chain.

SFG spectroscopy has established itself as a powerful nano-characterization method. The technique can probe any interfacial medium that is optically accessible, and has been used to probe many interfacial processes such as corrosion, surface phase transitions, detergency, and the structure of cell membranes.



**Figure 4.35** (a) A self-assembled dodecane thiol monolayer at the gold-water interface. An all-trans conformation due to strong inter-alkyl chain hydrophobic interactions creates a close-packed structure with the terminal  $\text{CH}_3$  groups pointing toward the water phase. The SFG spectrum in (b) shows only features due to the non-centrosymmetric  $\text{CH}_3$  groups. The  $\text{CH}_2$  groups are centrosymmetric and so do not appear in the SFG spectrum.



**Figure 4.36** (a) A self-assembled mercaptododecanoic acid monolayer at the gold-water interface. The large anionic headgroup prevents the monolayer from packing tightly and creates kinks and gauche defects in the chain. This randomized the headgroups. The SFG spectrum in (b) shows only features due to the non-centrosymmetric  $\text{CH}_2$  groups.

## 4.8 X-RAY SPECTROSCOPY

X-ray spectroscopy is a collection of techniques that use x-ray excitation in order to probe the electronic structure of molecules in a material. These techniques are particularly useful methods for determining a material's composition, probing the ordering of individual atoms in a crystal, or observing nanoscale molecular processes. Each technique is based on the measurement of absorption, emission, or scattering of electromagnetic radiation caused by x-rays.

### 4.8.1 Absorption

When x-rays pass through a thin layer of matter, the intensity of the x-rays is diminished as a result of absorption and scattering. The effect of scattering can be ignored in wavelength regions at which significant absorption occurs. Each element has its own absorption spectrum with well-defined x-ray absorption peaks that can be used to identify it. In the x-ray region of light, there is enough energy to cause observable changes in the electronic state of a molecule. When x-ray energy corresponding to the binding energy of core electrons is absorbed, the core electron is ejected from the atom. This results in an excited ion. There is a higher probability of this happening when an x-ray beam with energy equal to the binding energy of the core electron is used. As the x-ray beam increases in energy, away from the core electron binding energy, the probability of the corresponding wavelength being absorbed diminishes, and so the amount of the x-ray beam absorbed will decrease. If the x-ray beam is too low in energy, the corresponding wavelength will not be appreciably present and will not be able to eject the core electron. This will cause an abrupt decrease in the amount of the x-ray beam absorbed. The wavelength of x-rays ranges from about 10 to  $10^{-6}$  nm, but conventional x-ray spectroscopy generally uses only wavelengths between 2.5 and 0.001 nm because this range contains the x-rays with energies corresponding to core electron-binding energies, which differ between elements.

### 4.8.2 Fluorescence

The excited ion, which is a result of the x-rays' ejecting of a core electron, will fluoresce through transitions of electrons in higher energy levels to the vacancy left by the ejected core electron. These transitions allow the excited ion to return to its more stable ground state. This fluorescence is measurable and is often used in nanoscience techniques

such as fluorescence recovery after photobleaching (FRAP), fluorescence resonant energy transfer (FRET), and fluorescence interference contrast microscopy (FLIC), which are covered in the next chapter. X-ray fluorescence is also commonly used in dyes. These fluorescing dyes can be attached to target molecules of interest that can be excited and observed in conditions where dyes in the visual wavelengths of light are not usable. This is often the case for *in vivo* studies when there is a layer between the observer and the dyed molecule that allows for x-rays to pass through but does not allow visible light to pass through.

### 4.8.3 Diffraction

When x-rays pass through matter, the radiation interacts with electrons in the matter in such a way that the path of the x-rays can be altered. This scattering effect of matter on x-rays is known as diffraction. In a crystal, or any ordered sample, the x-rays scatter in ways that produce higher-intensity areas and lower-intensity areas, also termed constructive and destructive interference, respectively. By comparing these high- and low-intensity areas, it is possible to determine the architecture and the ordering in matter at the nanoscale. It is, of course, slightly more complicated in practice, because when an x-ray beam strikes an ordered crystal, which can be thought of as multiple layers of atoms, each subsequent layer of atoms scatters some of the beam and lets the remainder of the beam through. So that this diffraction can take place in a way that allows us to relate it to the materials structure, the space between layers of atoms in the material must be about the same distance as the wavelength of the radiation used to probe it, and the atoms comprising the system must be highly ordered with few defects.

In order to use the high- and low-intensity areas that result from shining x-rays through an ordered crystal, a way to relate them to the structure must be known. Fortunately, in 1912, W. L. Bragg determined this relationship. The Bragg equation for constructive interference is as follows:

$$2 d \sin \theta = n \lambda \quad (4.36)$$

In this equation  $d$  is the lattice spacing, or the distance between atoms in the subsequent layers.  $\theta$  is the incident angle between the beam of x-rays and the plane of the atom layer.  $\lambda$  is the wavelength of the x-rays. Lastly,  $n$  is an integer related to the order of the reflection.

## 4.9 IMAGING NANOSTRUCTURES

One of the more straightforward methods that we have not yet discussed is taking a “picture” of the nanostructure of interest. Indeed, so-called nano-imaging methods are often among the first used to study a specific nanomaterial. There are a variety of nano-imaging methods that exist. In the next few sections, we discuss a sampling of these methods, ranging from sub-diffraction limit optical methods such as near-field scanning microscopy, to other contrast mechanisms with the capability of resolving individual atoms such as the scanning probe microscopies.

### 4.9.1 Imaging Ellipsometry

#### 4.9.1.1 *Imaging Using Conventional Ellipsometry*

We have already encountered ellipsometry as a powerful method for determining the thickness of films as small as a few angstroms (for a review of ellipsometry, see Section 4.3). This method involves analyzing the change in the polarization state of a light beam upon reflection off a planar surface. The ellipsometric parameters recorded from monitoring this change in the polarization state of the light beam can be translated into thickness values for the surface after using an appropriate mathematical model. It should also be highlighted that a thickness value obtained in this manner actually represents the average thickness value within the beam spot on the surface. In order to achieve a better lateral (or horizontal) resolution, the light beam must be focused on a smaller spot on the surface. Thus, perhaps the most straightforward approach to image a surface using ellipsometry would be to maximize the lateral resolution by using a tightly focused light beam then determining the film thickness as the beam scans the surface. We could then obtain a three-dimensional image of the surface topography. However, using conventional ellipsometry to scan the sample in this way can be time consuming, so quicker methods called imaging ellipsometry have been developed.

#### 4.9.1.2 *Principles of Modern Imaging Ellipsometry*

Modern imaging ellipsometry can be thought of as combining an ellipsometer and a microscope. Rather than using a tightly focused light beam and scanning it across the sample, modern imaging ellipsometers generally employ a light beam with a large diameter (often on the order of millimeters) such that the entire sample is illuminated. The reflected image of the entire object is then focused onto a high-resolution CCD

camera using an objective lens. From this image, ellipsometric data of the sample (including thickness) can be extracted pixel by pixel, generating a surface topography map of the sample. Therefore, the ellipsometric data for the entire sample can be determined quickly, resulting in a 3-D image of the sample.

#### **4.9.1.3 Methods for Extracting Ellipsometric Data in Imaging Ellipsometry**

The “pixel-by-pixel” ellipsometric data from an imaging ellipsometer setup can be determined using one of several methods. A popular method is to employ “off-null” mode analysis. We recall from our discussion of ellipsometry in Section 4.4 that conventional ellipsometers are often operated in null mode. In null mode, the change in the polarization state of the reflected light is determined by changing the polarizer, compensator, and analyzer so that a null condition is achieved (or so that no light passes through the analyzer). Depending on the values of the polarizer, compensator, and analyzer used to achieve this null condition, different ellipsometric parameters may be calculated. In “off-null” mode, the imaging ellipsometer is “zeroed” by determining the null conditions for the bare substrate. These conditions are then kept constant as the entire sample is imaged. Since the sample is generally not of the same thickness and refractive index as the substrate, light of greater and greater intensity passes through the analyzer for thicker regions of the sample. Therefore, the intensity of the light in each pixel can be related to the thickness of the sample at that point. It should be noted, however, that this type of off-null mode analysis is substrate, sample, and thickness dependent, so comparisons with a reference material of known thickness are often recommended. For example, for a pure silicon substrate and a biological sample, the intensity  $I$  under an off-null condition has been reported to be related to the film thickness  $d$  by  $I = kd^2$  where  $k$  is a proportionality constant, and this equation is valid with a deviation of approximately 2% up to  $d \sim 5\text{nm}$ .

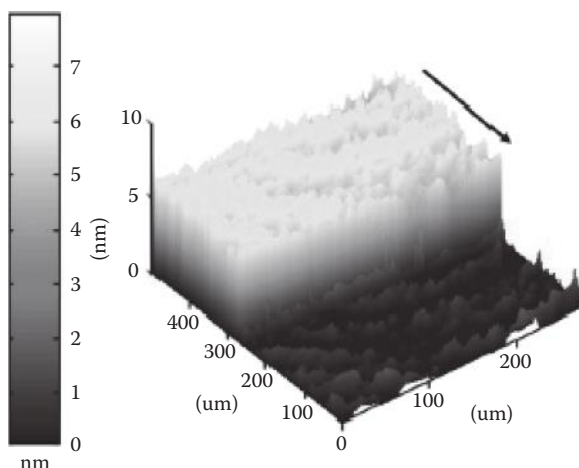
An alternative and perhaps more straightforward method for obtaining the “pixel-by-pixel” ellipsometric parameters is to continually adjust the polarizer, compensator, and analyzer in order to determine the null condition for each pixel (or group of pixels) and then extract the parameters from those null conditions as would be done in conventional ellipsometry. While this approach is slightly more time-consuming, it does not require the use of any reference samples or the intensity-thickness assumptions employed in “off-null” mode analysis.

#### 4.9.1.4 Image Focusing

One of the major obstacles in imaging ellipsometry is that only the center of the image is in focus, due to the large incident angle of the light beam. To compensate for this limited focus area, a scanner can be incorporated into imaging ellipsometer instruments. The image may then be taken at several different positions, and the computer software combines these images into one focused image of the entire sample. An alternative approach has been to use the Scheimpflug method, which is an optical technique that allows the entire sample to be placed in focus by tilting the lens to a certain degree, called the Scheimpflug line.

#### 4.9.1.5 Resolution of an Imaging Ellipsometer

The lateral ( $x,y$ ) resolution for an imaging ellipsometer as described above is generally limited by the resolving power of the CCD camera and is usually on the order of a few microns. The depth ( $z$ ) resolution, on the other hand, is comparable to conventional ellipsometry at a few tenths of a nanometer. Because of its nondestructive nature, compatibility with a wide variety of sample types, and excellent temporal and depth resolution, imaging ellipsometry is a powerful technique to image nanostructures, and is suited for thin film analysis. Figure 4.37 shows an ellipsometric thickness image of a single phospholipid bilayer assembled on a glass surface.



**Figure 4.37** An ellipsometric thickness image showing the height of a phospholipid bilayer assembled on a solid support. (Image provided by Professor Atul Parikh, University of California, Davis.)



The bilayer thickness is around 5 nm. The figure shows how imaging ellipsometry can be used to follow the slow hydration-induced spreading of the bilayer over the uncovered portion of the substrate. Thus, imaging ellipsometry provides topographical data with excellent vertical resolution. However, because the lateral resolution is on the order of microns, nanostructures cannot be resolved in the x-y plane. In order to image nanostructures at atomic-level resolutions in all three dimensions, methods such as the scanning probe microscopies must be used.

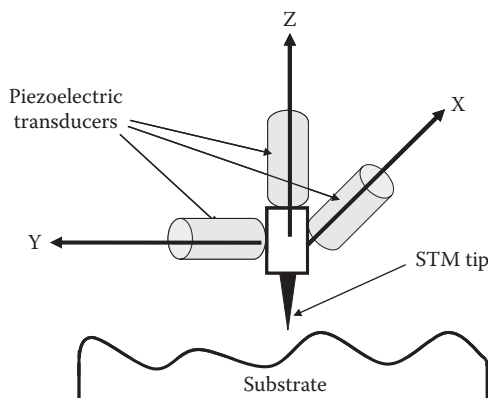
#### **4.9.2 Scanning Probe Methods**

The scanning probe microscopies are powerful imaging techniques in which a very sharp tip is scanned across a surface, producing an image with near atomic-level resolution. Scanning tunneling microscopy (STM) and atomic force microscopy (AFM) are the most common examples of scanning probe methods. Both STM and AFM produce three-dimensional images of surfaces that approach atomic or molecular resolution, making them ideal for the study of nanostructures on surfaces.

##### ***4.9.2.1 Scanning Tunneling Microscopy***

STM operates by monitoring the “tunneling” current that is produced when a sharp tip is brought extremely close to a surface that is able to conduct electricity. In order to bring and maintain the tip so close to the surface, a piezoelectric transducer is used. As a refresher, we recall that piezoelectric transducers possess the ability to physically expand or contract in response to an applied voltage (for a more complete discussion, see Section 4.3). Therefore, the sharp metal STM tip, usually made of platinum, is attached to a piezoelectric scanner that can cause the tip to move small distances in the x, y, and z directions. In fact, depending on the type of piezoelectric material used to make the scanner (usually some form of ceramic) and its dimensions, the contraction or expansion of the material can be as small as 1 nm per volt applied, allowing for the metal tip to be brought close to the surface being imaged. In early models of STMs, the piezoelectric scanners consisted of transducers arranged in the x, y, and z directions, as shown in Figure 4.38, but more recent models have made use of a tube-like piezoelectric transducer to achieve better resolution.

Using a piezoelectric scanner, the STM tip can be brought close (within 1 nm) to a conducting surface. A small voltage, generally between 2 mV and 2 V, is then applied between the



**Figure 4.38** Schematic diagram of the piezoelectric transducers that control the position of an STM tip near the sample surface. For many piezoelectric transducers used, a distance of as little as 1 nm can be affected with a single applied volt. When the STM tip is positioned sufficiently close to the sample surface and a potential is applied across them both, a tunneling current will be induced between the tip and the sample. This tunneling current is the basis of the STM measurement.

conducting substrate and the metal tip, causing electrons to tunnel between the tip and the surface, creating a current. The magnitude of this tunneling current depends exponentially on the distance ( $h$ ) between the conducting surface and the STM tip. For a constant applied voltage, the tunneling current,  $I$ , is given approximately by

$$I \approx e^{-2\kappa h} \quad (4.37)$$

Equation 4.37 tells us that the tunneling current decreases exponentially as we move the tip away from the surface. The constant  $\kappa$  is called the electronic decay length of the electron and is a measure of how the probability density of a confined electron decays with distance or, more precisely, it is the decay length of the electronic wavefunction. Not surprisingly,  $\kappa$  depends on how tightly the electron is bound to the conducting surface, called the *work function* of the surface. The ability of electrons to travel or “tunnel” through a non-conducting medium between two conducting materials that are close together is strictly a quantum mechanical effect. The details of this effect are not covered here. Suffice it to say that electrons can tunnel between two conducting materials, and this tunneling ability is exponentially dependent on the distance between the two materials.

As the STM tip is scanned across a rough surface, the tunneling current changes as the tip encounters bumps or dips in the surface (or as  $h$  changes). There are two obvious methods for imaging the surface. In the first method, known as “constant-height” STM, the tip is held at a constant vertical position and the tunneling current is plotted as a function of the x-y region scanned. Because the tunneling current is dependent on the distance between the tip and the surface ( $h$ ), the surface can be imaged. The second method, known as “constant-current” STM, is to maintain a constant tunneling current as the tip is scanned across the surface by varying the tip’s vertical position. The piezoelectric transducer moves the tip up or down as the surface is scanned, ensuring that the tunneling current is maintained at the predetermined value. Essentially,  $h$  remains constant throughout the scanning process, so the instrumentation has to monitor only the changes in the tip’s vertical position to provide a topographical image of the surface. In practice, this second method is generally used to generate a surface image.

As with most of the scanning probe microscopies, STM surface images are typically generated by scanning the tip across the surface in a *raster pattern*, or line by line, to create a rectangular grid. In some STM models, the scanning process is controlled by moving the substrate and not the tip, which is moved only vertically to maintain a constant tunneling current.

#### Example 4.8 Current Changes in STM

Consider a metal surface with an electronic decay length  $\kappa$  of  $10 \text{ nm}^{-1}$ . By how much does the tunneling current change when the height  $h$  between the surface and the STM tip increases from  $1.0 \text{ nm}$  to  $1.1 \text{ nm}$ ?

*Solution* From Equation 4.37, we have  $I \approx \exp(-2\kappa h)$ . We can determine the approximate factor by which the tunneling current changes by taking the following ratio:

$$\frac{I_{1.0\text{nm}}}{I_{1.1\text{nm}}} \approx \frac{\exp(-2 \times 10 \text{ nm}^{-1} \times 1.0 \text{ nm})}{\exp(-2 \times 10 \text{ nm}^{-1} \times 1.1 \text{ nm})} \approx 8$$

Thus, we see that the current changes by almost an order of magnitude when the distance  $h$  varies by  $0.1 \text{ nm}$ .

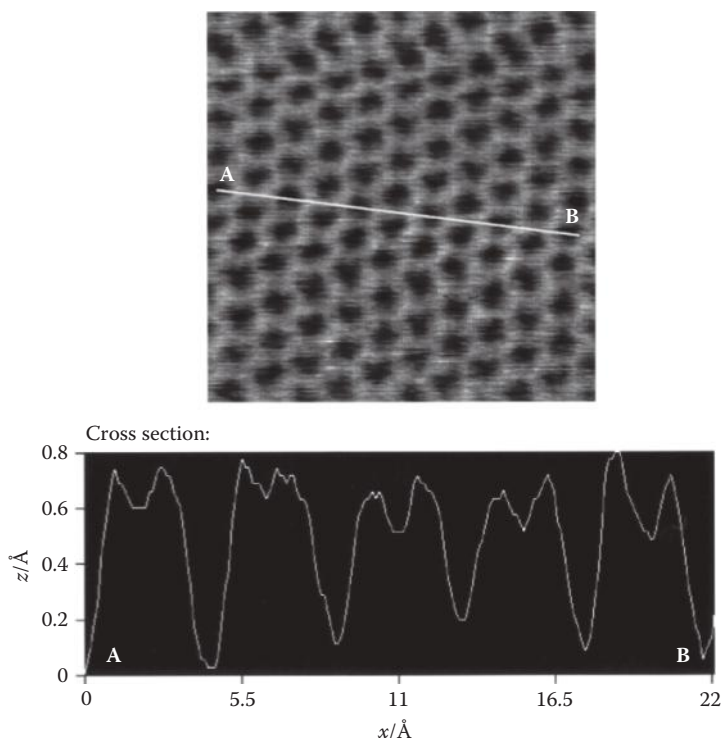
Example 4.8 emphasizes the sensitivity of STM to changes in height. In fact, if the tunneling current is kept constant then

the height remains constant to within  $10^{-3}$  nm, meaning that atomic-level resolution is achieved.

While such resolutions constitute a major achievement, STM suffers from its inability to image surfaces that do not conduct electricity. Non-conducting surfaces are unable to be imaged unless they are modified or coated with a conducting substance. Atomic force microscopy, discussed below, does not suffer from this disadvantage, although STM provides superior resolution.

For a typical STM, the lateral range of the scanner is generally from tens of angstroms to approximately  $100\text{ }\mu\text{m}$ . The allowable height is from the sub-angstrom to  $\sim 10\text{ }\mu\text{m}$  range. The tip is generally constructed by manually cutting a platinum wire or by electrochemical etching of tungsten metal.

One example of an application of STM is the atomic resolution STM image of a graphite surface, shown in Figure 4.39. In this image, the hexagonal arrangement of the carbon atoms



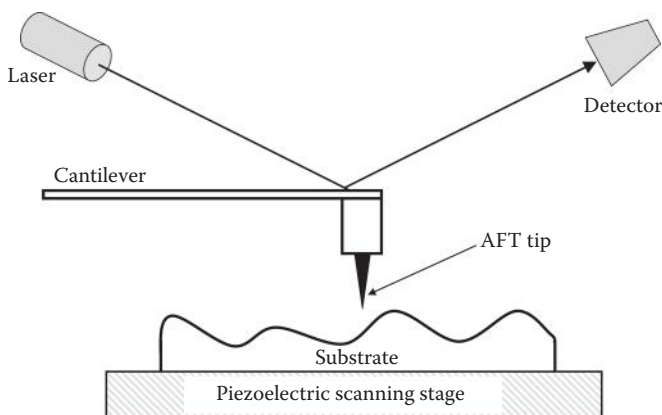
**Figure 4.39** STM image of a graphite surface. Notice the angstrom-level resolution that is characteristic of STM. (Image from Atamny et al, *Phys. Chem. Chem. Phys.*, 1999, 1, 4113–4118. Reproduced by permission of the PCCP Owner Societies.)

in graphite is clearly discernible, as well as the valleys in between adjacent carbon atoms, thus demonstrating the utility of STM in studying conducting surfaces.

#### 4.9.2.2 Atomic Force Microscopy

Atomic force microscopy, or AFM, operates on principles similar to those of STM. As with STM, AFM functions by scanning a sharp tip across a surface to generate an image. However, rather than monitoring the tunneling current between the tip and the surface, AFM monitors the height of the tip as it physically interacts with the surface at a constant force.

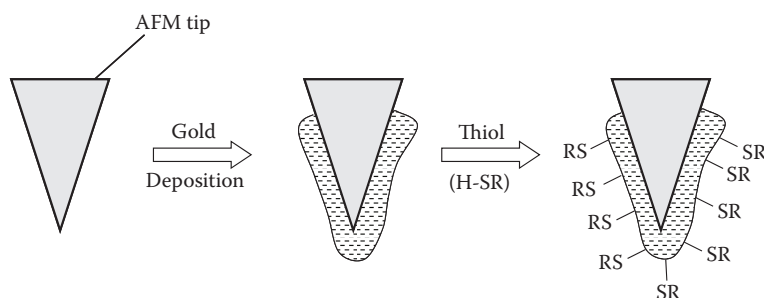
The setup of a typical AFM is shown in Figure 4.40. A sharp tip, often made of diamond or silicon nitride, is attached to a cantilever spring and placed in physical contact with the substrate at constant force. The position of the tip is monitored by reflecting a laser beam off the back of the cantilever and monitoring the deflection of this beam as the tip is scanned across the surface. If the tip encounters a bump or a dip, the reflection of the laser beam is slightly altered and monitored by the detector. Rather than moving the tip to scan the surface, the substrate is typically mounted on a piezoelectric tube scanner, allowing for the surface to be moved in a raster pattern under the tip. The cantilever itself is tens of microns in length, <10 microns in width, and ~1 micron in thickness. The tip is usually cone or pyramid shaped, with a height of several microns and a base width of several microns.



**Figure 4.40** Schematic diagram of an atomic force microscope. The AFM tip is held in constant force against the sample surface by the cantilever. A laser beam is reflected off the back of the cantilever to monitor the tip height. The sample is scanned underneath the AFM tip by a piezoelectric stage.

AFM is able to image both non-conducting and conducting surfaces, unlike STM. However, in AFM, the physical contact between the tip and the surface can damage the surface, resulting in a distorted image. This drawback is particularly problematic if one wishes to image a “soft” surface such as a biological membrane or a surfactant film. Fortunately, this problem of surface destruction can be somewhat overcome by scanning the substrate in *tapping mode*. In tapping-mode AFM, the tip is “tapped” on the surface, placing it in contact with the surface only for a short amount of time. This “tapping” is usually accomplished by oscillating the cantilever with constant driving force and ensuring that the cantilever is positioned so that the tip touches the surface only at the bottom of the oscillation. Typical frequencies of oscillation are on the order of a few hundred kilohertz. By using tapping-mode AFM, surfaces that would normally be destroyed are able to be imaged properly.

AFM can be used under water as well as in air, allowing for imaging at the surface-liquid interface. This ability is particularly important for biological samples, which might be distorted at the surface-air boundary. One application of AFM in surface imaging is the detection of different functional groups on a surface, called chemical force microscopy (CFM). In CFM, the AFM tip is chemically functionalized, often by coating the tip with gold then functionalizing it with a self-assembled monolayer using thiol-gold chemistry (Figure 4.41). For example, a gold-coated AFM tip might be functionalized with 11-mercaptopundecanoic acid, effectively resulting in a  $-\text{COOH}$  coated AFM tip [R is a  $-(\text{CH}_2)_n\text{COOH}$  group in Figure 4.41]. This type of chemisorption is discussed in Chapter 5. This chemically

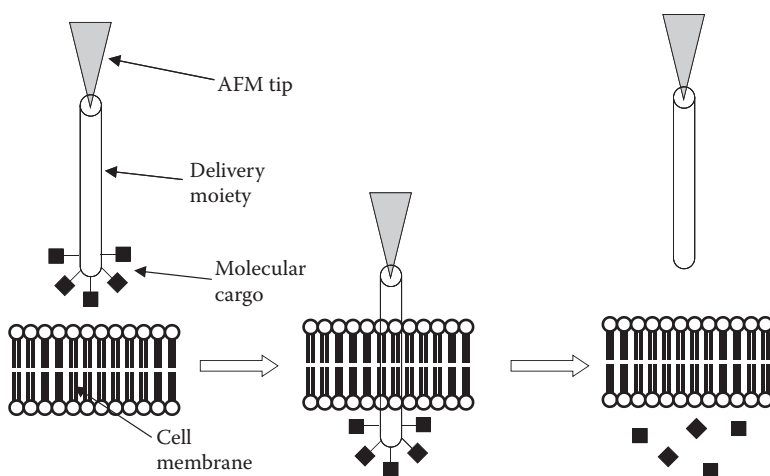


**Figure 4.41** A method for the creation of a CFM tip. An AFM tip is coated with Au and then is functionalized using thiol-gold chemistry. The functional group R can then be used to probe the surface.

functionalized AFM tip can then be scanned across a surface composed of different functional groups. Depending on the interaction between the surface-bound functional group and the functional group on the AFM tip, a different “height” is reported. A  $-\text{COOH}$  coated tip would be expected to experience more frictional force from surface-bound  $-\text{COOH}$  groups as compared to  $-\text{CH}_3$  groups, and therefore the  $-\text{COOH}$  groups would appear to be “taller” than the  $-\text{CH}_3$  groups. For example, a  $-\text{COOH}$ -coated AFM tip can be used to scan across a surface containing both  $-\text{CH}_3$  and  $-\text{COOH}$  groups. A different image is obtained when the same surface is scanned with a  $-\text{CH}_3$ -coated AFM tip. The results are striking and highlight the utility of CFM to detect different functional groups on a surface.

Aside from being used to image surfaces and surface-bound material, AFM has other applications. One unorthodox application of AFM in a non-imaging setting is in the recent development of a cell nanoinjector. In the past, in order to introduce a substance into a living cell, the cell had to be permeabilized with an electric current or chemical agent, or a bulky micropipette had to be used as an injector. This process could result in damage to the cell membrane, inducing undesirable side effects. However, scientists have recently been able to attach a delivery moiety (such as a carbon nanotube) to the end of an AFM tip for use as a “nanoinjector.” Because of the tip’s small dimensions, the cell membrane is not perturbed.

The substance (or cargo) to be injected into a cell is chemically attached to the delivery moiety (through a disulfide linkage, for instance). A carbon nanotube as a delivery moiety acts like a “nanoneedle” when it is attached to the AFM tip, which is then used to inject the carbon nanotube (with the molecular cargo) through the cell membrane. Inside the reducing environment of the cell interior, the disulfide linkage between the cargo and the carbon nanotube is severed and the cargo is released into the cell. The AFM tip then withdraws the nanoneedle from the cell interior. Because the nanoneedle is of such small dimensions, the cell membrane is not perturbed greatly and the cell remains unharmed. Furthermore, by using the AFM machinery, excellent spatial precision in the placement of the cargo can be achieved. Figure 4.42 shows a schematic diagram of a nanoinjector and Figure 4.43 shows electron microscope images of the nanoneedle before and after the attachment of molecular cargo. Although the AFM-operated nanoinjector is admittedly limited to delivering cargo to one



**Figure 4.42** Schematic steps of the AFM-operated nanoinjector. The nanoinjector is linked to the molecular cargo through a disulfide bond and then poised above the cell membrane. The nanoinjector is then inserted through the cell membrane and the molecular cargo is released inside the cell as the disulfide linkages are reduced. Finally the nanoinjector is withdrawn from the cell, ideally causing little to no membrane damage.

cell at a time, it still promises to be a useful tool in a wide variety of biological studies.

### 4.9.3 Transmission Electron Microscopy

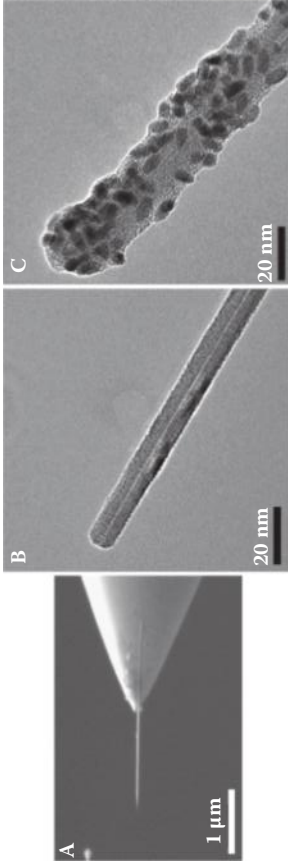
Unlike the scanning probe microscopies, which utilize physical interactions between a sharp tip and a surface to create an image, the electron microscopies use electrons rather than a beam of light to visualize a sample. Because electrons have much smaller wavelengths than visible light, electron microscopies are able to achieve resolutions far greater than optical microscopies.

Scanning electron microscopy (SEM) is a commonly used electron microscopy method that functions by monitoring the electrons that are backscattered after interacting with a sample and uses those backscattered electrons to reconstruct an image of the sample. On the other hand, transmission electron microscopy (TEM) makes use of electrons that are transmitted through a thin sample to produce an image. For our current discussion, we focus on TEM as a tool to image nanomaterials.

#### 4.9.3.1 Principles of TEM

The resolution of an image produced by an optical microscope is limited by the wavelength of the radiation being used. If





**Figure 4.43** Imaging of the nanoneedle before and after the molecular cargo is attached. (a) shows an SEM image of the carbon nanotube affixed to the AFM tip. (b) is a TEM image of the carbon nanotube shown in (a). (c) shows a TEM image of the nanoneedle after the molecular cargo has been attached (a quantum dot in this case). (Image taken from Chen et al, 2007, *PNAS*, 104, 8218–8222. Copyright 2007 National Academy of Sciences, USA. With permission.)

we use the Abbe equation applied to the resolution of a light microscope, then the maximum resolution  $\delta$  is approximately given as

$$\delta \approx 0.61 \left( \frac{\lambda}{n \sin(\beta)} \right) \quad (4.38)$$

where  $\lambda$  is the wavelength of the light being used,  $n$  is the refractive index of the viewing medium, and  $\beta$  is a property of the magnifying lens called the semi-angle of collection. Together,  $n \sin(\beta)$  is often called the numerical aperture (NA) of the objective. To provide a rough estimate of  $\delta$ , let's assume  $n \sin(\beta) = 1$ –1.5. We see that the resolution of a light microscope is approximately 50–60% of the wavelength of the light being used. Visible light has wavelengths in the range of ~350–750 nm, so conventional optical microscopies are unable to resolve objects that are smaller than a few hundred nanometers. Since most nanomaterials of interest possess structures that are much smaller than several hundred nanometers, optical microscopies are only moderately useful in the imaging of nanomaterials. In order to visualize these nanomaterials, something with a much smaller wavelength than visible light must be used.

Electrons, as all small particles, can be thought of as being both particles and waves (with characteristic wavelengths). Indeed, one of the great advances in early quantum mechanics was the development of the concept of wave-particle duality—that matter and light can be treated simultaneously as both particles and waves. De Broglie's famous equation relates the wavelength  $\lambda$  of a particle to its momentum  $p$  by

$$\lambda = \frac{h}{p} \quad (4.39)$$

where  $h$  is Planck's constant, equal to  $6.626 \times 10^{-34}$  J s. Therefore, in order to calculate the wavelength of an electron (which gives us an idea of the maximum possible resolution of an electron microscope), we must be able to calculate its momentum. This goal can be achieved by understanding the basic physics of a TEM.

In an electron microscope, an electron produced in the electron gun is accelerated by an electric potential  $V$  toward the sample to be studied. This acceleration imparts a kinetic energy

$K$  to the electron that is equal to the accelerating potential or voltage. Therefore, using classical physics, we can write

$$V = K = \frac{1}{2} m_e v^2 \quad (4.40)$$

where  $m_e$  is the rest mass of the electron ( $9.11 \times 10^{-31}$  kg) and  $v$  is the electron's velocity. From classical physics we also know that  $p = mv$ , so using Equation 4.40 we can write

$$p = m_e v = \sqrt{2m_e \left( \frac{1}{2} m_e v^2 \right)} = \sqrt{2m_e V} \quad (4.41)$$

This equation, then, shows that we can calculate the momentum of an electron if we know the accelerating voltage used in the TEM. Plugging this result into de Broglie's relationship, we get

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2m_e V}} \quad (4.42)$$

Therefore, we can calculate the wavelength of an electron from its accelerating voltage in an electron microscope. Furthermore, we see that as we increase the accelerating voltage, we decrease the wavelength of the electron. Ignoring other effects, an electron microscope can achieve better resolution by accelerating the electrons to higher energies. However, as the energy of the electrons increases, so does the likelihood that it can destroy or damage the sample being studied. This is one limitation to the maximum resolution that can be achieved in a TEM.

The magnitude of accelerating voltages typically used in most TEMs (on the order of hundreds of keV) accelerates the electron so much (near the speed of light) that relativistic effects must be accounted for. To account for these relativistic effects, Equation 4.42 becomes

$$\lambda = \frac{h}{\left[ 2m_e V \left( 1 + \frac{V}{2m_e c^2} \right) \right]^{1/2}} \quad (4.43)$$

where  $c$  is the speed of light in a vacuum. Using this equation we can calculate that an electron that has been accelerated to 100 keV has a relativistic wavelength of  $\sim 3.7$  pm. For a TEM, we can approximate the maximum resolution by modifying Equation 4.43 to

$$\delta \approx 0.61 \left( \frac{\lambda}{\beta} \right) \quad (4.44)$$

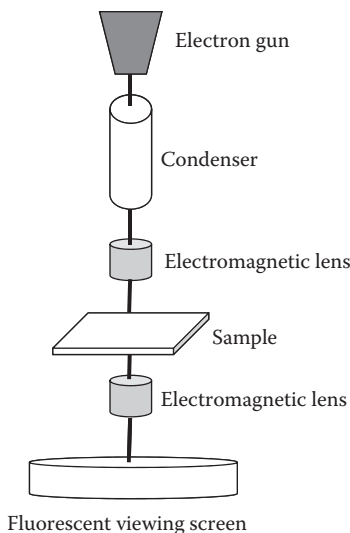
where  $\beta$  in this case is a property of the electron “lens” used to focus the electron beam. Therefore, using Equation 4.44 we see that with a 100-keV electron beam ( $\lambda \sim 4$  pm), we can achieve a theoretical maximum resolution ( $\delta \sim$  several picometers) that is smaller than the diameter of an atom. While such maximum resolution is not feasible because we are unable to build perfect electron lenses, it highlights the fact that TEM is able to achieve a high level of resolution.

#### 4.9.3.2 TEM Instrumentation

In a simplified way, a transmission electron microscope operates like a slide projector. In a slide projector a beam of light is transmitted through a slide. Some of the light is reflected or absorbed by the slide, so when the transmitted light is projected onto a screen, an image is produced. A TEM functions in essentially the same manner, with the only difference being that a beam of electrons is transmitted through the sample rather than a beam of visible light.

As seen in Figure 4.44, an electron gun produces a beam of electrons, often by heating a metal (most commonly tungsten) to such high temperatures that electrons are ejected from its surface. These ejected electrons are condensed and then focused by electromagnetic lenses, which are analogous to optical lenses, but function on different principles. The focused beam of electrons is shone on the sample, and those electrons that have been transmitted are gathered by another lens and then shone onto a fluorescent screen or other detector to produce an image of the sample. The setup of a TEM can be modified so that it is able to scan an entire sample, and this technique is called scanning transmission electron microscopy (STEM). As with AFM and STM, STEM typically scans a given sample in a raster pattern to produce an image.

In order to prevent deflection of the beam of electrons due to interactions with gas molecules inside the machine, the interior of a TEM must be operated under high vacuum conditions.



**Figure 4.44** Schematic diagram of a transmission electron microscope. Electrons are produced by an electron gun and condensed into an electron beam in the condenser. This electron beam is focused onto the very thin sample by electromagnetic lenses. The transmitted electrons are collected by another electromagnetic lens and projected onto a fluorescent screen where they produce a visible image, which can be viewed directly or monitored on a computer.

This requirement is one of the principal disadvantages of TEM, although recently there have been rapid developments in the field of environmental TEM that allow operation under lesser vacuums. We will focus on the more common, high vacuum TEM. For this technique, the sample being examined must be able to withstand the high vacuum; otherwise, the image won't be an accurate representation of the sample under normal conditions. For non-solid samples, two principal methods of sample preparation have emerged. These methods allow for the imaging of biological samples or other soft nanomaterials that otherwise could not be imaged using TEM. The first method is to dehydrate the sample, then stain the sample or coat it with metal to produce the necessary contrast. The second is to cryogenically freeze the sample and image the frozen specimen. One requirement of this second method is that the sample must be frozen so quickly that it does not have time to rearrange into a crystalline state. Otherwise, the ordered ice crystals produce a diffraction pattern of the electron beam that obscures the image of the sample. If the sample is frozen quickly enough, amorphous ice is produced and one can obtain an accurate image of the sample in its “natural” state.

The second major difficulty with TEM is the thickness of the sample. In order to be effectively imaged, the sample must be thin enough to be transparent to the incoming beam of electrons. As a general rule, the sample should be less than 100 nm thick, although the exact suitable thickness depends on the material being examined and the energy of the electrons used. Thicker samples can be examined if electrons with higher energies are employed, but at higher energies the electrons may begin to destroy the sample. In terms of sample mounting, the sample can sometimes be placed on a sample holder that is very thin and transparent (or mostly transparent) to the electron beam. For extremely small samples, a sheet of thin, amorphous carbon films have been used as sample holders. These thin carbon films are suitable for such applications because they are electron-transparent nanomaterials that possess considerable durability even at a thickness of only one atom.

There are difficulties associated with using TEM to image nanomaterials, but it is a useful technique for a variety of applications and boasts impressive resolutions. Normally TEM is recommended for use in conjunction with other imaging methods to obtain an accurate understanding of the material being studied.

Figure 4.43 shows a TEM image of a carbon nanotube that has been attached to an AFM tip to serve as a nanoinjector (the operation of a nanoinjector is discussed in the previous section). The molecular cargo attached to the surface of the nanotube is clearly visible, thus demonstrating the resolving power of TEM.

#### 4.9.4 Near-Field Scanning Optical Microscopy

In our introduction to TEM, we discussed the limited utility of conventional light microscopies in imaging nanostructures. In general, the resolution of an optical microscope is limited by the size of the spot to which the light beam can be focused using magnifying lenses. This limitation is often referred to as the *diffraction limit*. The diffraction limit on resolution is wavelength dependent, as given by the Abbe equation

$$\delta \approx 0.61 \left( \frac{\lambda}{NA} \right) \quad (4.45)$$

where  $\delta$  is the resolution of the microscope,  $\lambda$  is the wavelength of the light, and  $NA$  is the numerical aperture of the objective.

For modern objectives examining samples in an aqueous medium,  $NA$  is usually in the range of 1.3 to 1.5. Therefore, the resolving power of a conventional optical microscope is approximately half the wavelength of the incident light, typically  $\sim 200$  nm for visible light.

Because of the diffraction limit on optical microscopies, alternative methods have been developed to examine objects that are smaller than several hundred nanometers. The electron microscopies (TEM and SEM) and the scanning probe microscopies (STM and AFM) discussed earlier are the prototypical examples. However, these alternative methods generally suffer from limited applicability to only specific types of samples. For example, the electron microscopies require samples that are resistant to very high vacuums and STM requires samples that are conductive. Furthermore, none of these alternative methods offer the types of information that are available to optical methods—such as spectroscopic information, excellent time resolution, fluorescence detection capabilities, information about refractive index and reflectance of the sample, and contrasting power using different staining agents.

Near-field scanning optical microscopy (NSOM or SNOM) is an optical microscopy that can operate with resolutions below the diffraction limit, and therefore is able to offer the advantages of optical microscopies with a resolution that is actually useful for the study of nanomaterials.

#### ***4.9.4.1 History and Principles of NSOM***

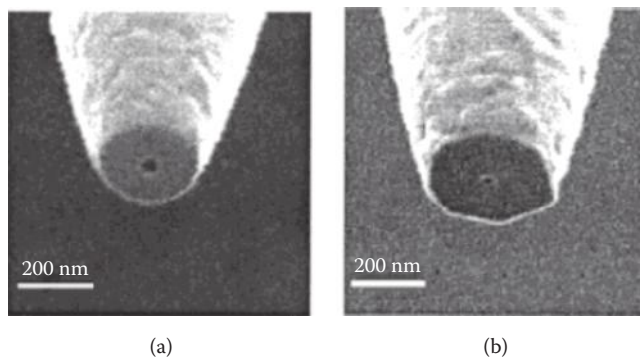
Although the fundamental idea of NSOM is relatively simple, its practical implementation proved to be rather difficult. The original idea was developed by Edward Synge and was published in a series of papers beginning in 1928. Synge realized that if the diffraction limit was imposed by the practical limit to which a beam of light could be focused then the limit might be overcome by shining light through a very small hole (or aperture) that was smaller than the wavelength of the light itself. If this hole were placed close to the sample, the light would not have time to diffract outward and destroy the resolution of the image. Thus, a sample could be imaged at a resolution below the diffraction limit.

Despite Synge's development of the theory of NSOM in the early 20th century, it was not implemented until 1972 when Ash and Nicholls used an NSOM setup with microwave radiation ( $\lambda \sim 3$  cm) to image a metal grating sample. They demonstrated that a resolution of 1/60th of the wavelength of the

incident radiation was achievable using their method. Their results validated Synge's theory, but practical considerations prevented the development of an NSOM using the much smaller wavelengths of visible light until the mid-1980s. During this period, scientists overcame the technical difficulties of implementing NSOM with visible light. Their practical setup serves as the basis for modern NSOM instruments.

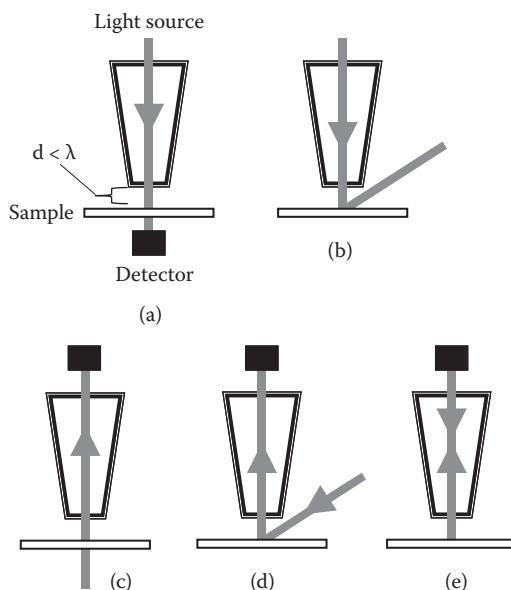
#### 4.9.4.2 Modern NSOM Instrumentation and Different NSOM Operating Modes

The major component of any NSOM microscope is the aperture tip or NSOM probe. A variety of NSOM probes exist. One common NSOM probe is manufactured by heating and pulling a fiber-optic cable into a very fine point, then coating the tapered end with reflective metal, except for a very small aperture at the point. SEM images of this type of NSOM probe are shown in Figure 4.45. Laser light shone through the fiber-optic cable emerges from the aperture as a beam with a diameter that is smaller than the wavelength of light. For this type of fiber-optic NSOM probe, the fundamental maximum resolution is  $\sim 12$  nm, but the practical limit is typically  $\sim 50$  nm. Another common NSOM probe can be made by using electron beam lithography to create a nanometer-scaled aperture through silicon or another metal and then shining the laser



**Figure 4.45** SEM images of two NSOM probes that have been constructed by tapering a fiber-optic cable into a very fine point and then coating that tip with aluminum, leaving only a very small aperture at the point. The ends have also been flattened using a focused ion beam. In these NSOM probes the apertures are approximately (a) 120 nm and (b) 35 nm. (Image reprinted with permission from Veerman et al., 1998, *Appl. Phys. Lett.*, 72, 3115–3117 as shown in Dunn, *Chem Rev*, 1999, 99, 2891–2927. Copyright 1998 American Institute of Physics.)





**Figure 4.46** The different operating modes of NSOM. (a) In Transmission mode, the light travels from the probe through the sample to a detector on the other side. (b) In Reflection mode the light from the probe is reflected off the sample surface and captured by the detector. (c) In Transmission-Collection mode, the sample is illuminated from underneath and captured by the probe, through which the light travels to a detector. (d) In Reflection-Collection mode, the sample is illuminated externally to the probe and the reflected light is captured by the probe and channeled to the detector. (e) In Illumination-Collection mode, the probe is responsible for both illuminating the sample and collecting the light that reflects from its surface.

through this nanometric aperture. Both of these probes can be manufactured with aperture diameters of  $\sim 50\text{--}100\text{ nm}$ . Recent developments have also been made in so-called apertureless NSOM probes. These probes exhibit great promise in achieving even better sub-diffraction limit resolutions.

An NSOM setup can be operated in at least five different modes, as listed below (see Figure 4.46).

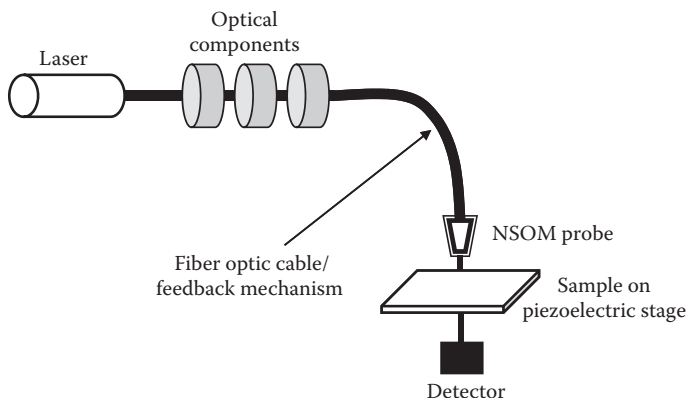
1. Transmission mode. Light is shone through the NSOM probe then detected on the other side of the sample.
2. Reflection mode. Light is passed through the NSOM probe and reflected from the sample. This reflected light is detected and compiled to produce an image.
3. Transmission-collection mode. Light is shone on the sample from underneath and the transmitted light is collected in the NSOM probe and passed to a detector.

4. Reflection-collection mode. Light is shone on the sample from above and reflected light is collected by the NSOM probe and sent to a detector.
5. Illumination-collection mode. The NSOM probe is used as both the light source and the light collector.

Depending on the type of sample examined and the kind of study undertaken, different NSOM operating modes may be employed.

Aside from issues related to manufacturing a suitable NSOM probe, the second major difficulty in the development of NSOM was probe positioning. In order to function correctly, the NSOM probe must be placed close to the surface of the sample being studied, usually less than a few nanometers. Such precise positioning is no easy task, especially when the probe is scanned across a “rough” sample surface. Furthermore, if the tip comes into forceful contact with the surface, either the probe or the sample could be damaged. Therefore, scientists have developed different feedback mechanisms to ensure that the probe remains at the correct distance above the sample. The shear-force feedback mechanism and the tapping-mode feedback mechanism are among the most common feedback mechanisms employed. While the details of each feedback mechanism are beyond the scope of this text, the essential idea behind each mechanism is that the probe is oscillated at a certain frequency and the force that is placed on the probe as it approaches the surface is monitored. By using a continuous feedback loop, the probe can be oscillated at the correct height above the sample. Furthermore, if the tip position is monitored as the sample is scanned in the x-y plane, then topographical information about the sample can be obtained (much like with AFM or STM) in addition to any optical information. Finally, to ensure accurate probe positioning, the entire NSOM instrument is usually operated on a vibration isolation table.

A schematic of a typical modern NSOM microscope is shown in Figure 4.47. The setup depicted is an NSOM microscope with a fiber-optic probe operating in transmission mode. The laser light is initially passed through a bandpass filter and a combination of half-wave and quarter-wave plates to control the wavelength and polarization of the light, respectively. The light then enters the fiber-optic cable and is transmitted to the probe tip through which it is shone on the sample. The transmitted light is collected by the detector and processed into an image. The height of the probe tip is controlled by the



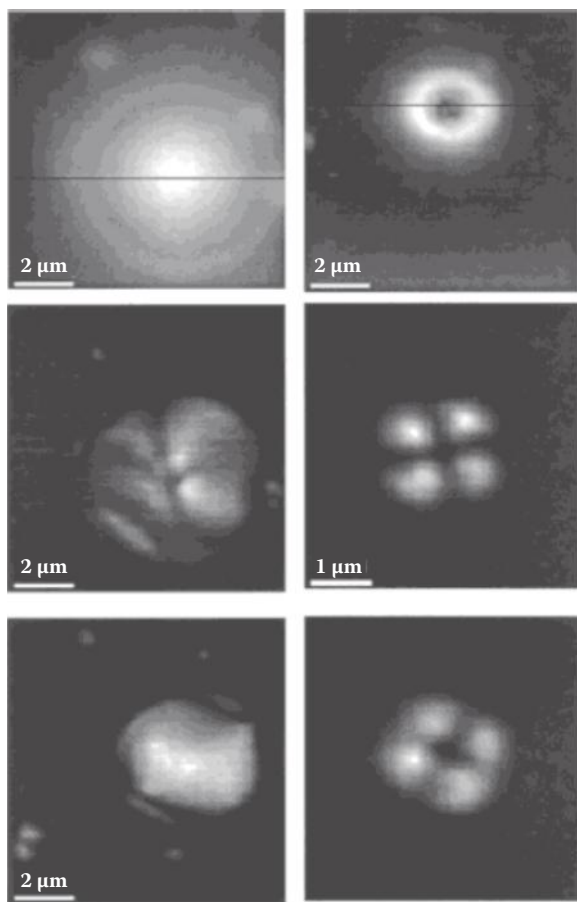
**Figure 4.47** Schematic of a near field scanning optical microscope operating in Transmission mode. Laser light is channeled into a fiber-optic cable that ends in an NSOM probe tip positioned very close to the sample surface. The transmitted light is captured by the detector to produce an image. The sample is scanned by the piezoelectric scanning stage, and the NSOM probe height is monitored and controlled by the feedback mechanism.

feedback mechanism, and the probe's vertical position can be monitored to produce topographical information. The sample itself is placed on a piezoelectric stage that can move in the x-y direction, allowing for the sample to be scanned in a raster pattern, much like with AFM or STM.

An example of an NSOM image is shown in Figure 4.48, which shows the topographical and cross-polarized optical NSOM images of spherical and toroidal (donut-shaped) liquid crystal droplets suspended in a polymer. Liquid crystals are materials that have properties between those of a liquid and those of a solid. There are many types of liquid crystals based on the orientation of molecules, but for our present purposes we merely note that the general shape of the droplets is clearly seen in the topographical images and that the orientation of the crystals can be seen in the cross-polarized images.

NSOM setups are often used in conjunction with other surface imaging methods to provide complementary information. For example, NSOM microscopes are commonly incorporated into conventional fluorescence microscopes, and surface-enhanced Raman NSOM setups have also been employed.

Overall, NSOM is a useful imaging method that can offer a wealth of information about a nanomaterial being studied. While its maximum resolution (often  $\sim 50$  nm) is not quite as impressive as some of the other imaging methods we have discussed, its versatility in being able to image a wide variety of samples under “normal” conditions and its richness in the



**Figure 4.48** NSOM images of spherical and toroidal liquid crystals suspended in a polymer. (Image from Mei et al, *Langmuir*, 1998, 14, 1945–1950 as shown in Dunn, *Chem Rev*, 1999, 99, 2891–2927. With permission.)

types of information it provides usually compensate for this lower resolution.

#### 4.10 LIGHT SCATTERING METHODS

Surface tension measurements can be used to study the formation of a surfactant monolayer at the air-water interface. Beyond the critical micelle concentration (CMC), surface tension values from such solutions become constant, indicating the formation of colloidal particles such as micellar phases or nanoparticles, which have strong tendencies to scatter radiation, especially x-rays or a beam of neutrons. There are three

ways that particles can scatter light in which there is no energy transfer (elastic light scattering): Rayleigh scattering, Mie scattering, and geometric scattering. The type of scattering depends on the size of the particle interacting with electromagnetic radiation. Consider the unitless parameter,  $\alpha$ , defined as,

$$\alpha = \frac{\pi D}{\lambda} \quad (4.46)$$

where  $\pi D$  is the circumference of the particle and  $\lambda$  is the wavelength of the incident light. Rayleigh scattering occurs when  $\alpha \ll 1$ , or when the particles are small compared to the wavelength of light. Mie scattering occurs when the particles are the same size as the wavelength of light (i.e.,  $\alpha \approx 1$ ) and geometric scattering occurs when the particles are relatively large (i.e.,  $\alpha \gg 1$ ).

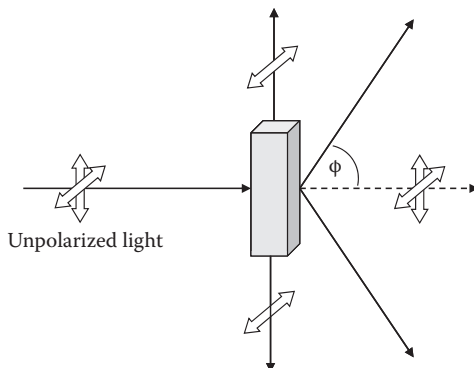
#### 4.10.1 The Measurement of Scattered Light:

##### Determining the Aggregation Number of Micelles

The intensity, angular distribution, and polarization of the scattered light depend on factors such as the shape and size of the particles, as well as on the interactions between them. Light scattering experiments can thus provide useful structural information (particle shape and size) and interparticle interactions on colloidal systems. The measurements are generally instantaneous, non-invasive, and allow representative sampling of polydispersed samples. However, the presence of small particle impurities, particularly those that have a tendency to scatter the light, can impose serious errors in the measurement.

In practice, a collimated beam of light of a given wavelength ( $\lambda$ ) and intensity ( $I_o$ ) passes through a solution containing the dispersed nanoparticles (Figure 4.49). The intensity of the scattered light is then measured as a function of the angle ( $\phi$ ) between the incident beam and the scattered beam.

We will begin by discussing how light scattering can be used to determine the aggregation number of a micelle. Micelles have diameters typically on the order of a few nanometers. The wavelength of visible light is about two orders of magnitude greater. Let's consider a beam of visible light passing through an aqueous solution containing spherical micelles. The solutions can be described in terms of two refractive indices, the refractive index of the randomly dispersed nano-spherical micelles ( $n_{\text{micelle}}$ ) and the refractive index of the continuous solvent ( $n_{\text{solvent}}$ ). These two refractive indices have different



**Figure 4.49** The scattering of unpolarized light through a sample. The intensity of the scattered light is measured as a function of the angle ( $\phi$ ) between the incident beam and the scattered beam.

values, and the average refractive index of the solution will therefore vary with the local concentration of micelles. This variation will cause light to be scattered. The intensity of the scattered light depends on the intensity and wavelength of the incoming light, the solution refractive index increment (i.e., how  $n$  varies with concentration,  $dn/dc$ ), and the average number density of micelles in the solution ( $N$ ). These parameters can be determined experimentally and are used to obtain the optical constant  $K_o$  (Equation 4.47).

$$K_o = 2\pi^2 \frac{n^2}{\lambda^4 N} \left( \frac{dn}{dc} \right)^2 \quad (4.47)$$

In order to understand how the intensity of light varies with the scattering angle  $\phi$ , we need to describe the incoming unpolarized light as being composed of two mutually perpendicular polarized components (Figure 4.49). For small scattering angles ( $\phi \sim 0$ ), these components will contribute equally to the scattered intensity. At very large scattering angles, one of the two polarized components contributes to a greater degree to the scattering. In fact, when  $\phi = 90^\circ$  the component polarized along the direction of the scattered beam has no contribution to the scattered intensity. By measuring the intensity of scattered light as a function of  $\phi$ , we can determine a quantity known as the Rayleigh ratio (Equation 4.48).

$$R_\phi = \frac{d^2}{1 + \cos^2 \phi} \frac{I}{I_o} \quad (4.48)$$

In this equation,  $d$  is the distance between the sample and the detector. The Rayleigh ratio will be different for a micellar solution compared to the pure solvent. In fact, it can be shown that the difference between these two Rayleigh ratios is given by Equation 4.49.

$$\Delta R_\phi = \frac{2\pi^2 n^2}{\lambda^4} \left( \frac{dn}{dc} \right)^2 RTc \sqrt{\frac{d\Psi}{dc}} \quad (4.49)$$

The term  $RTc\sqrt{d\Psi/dc}$  is known as the concentration fluctuation factor and describes the free energy cost in creating an inhomogeneity in micelle concentration. For micellar systems, the concentration fluctuation factor causes a dramatic change in the scattering intensity, as we go from monomers to predominantly micelles. Therefore, one would expect that  $\Delta R_\phi$  would be related to both the CMC of the surfactant and the concentration of the surfactant monomers ( $c_m$ ). The exact dependence is shown in Equation 4.50.

$$\Delta R_\phi = \frac{K_o(c_m - \text{CMC})}{10^3/M + 2B(c_m - \text{CMC})} \quad (4.50)$$

$M$  represents the molecular weight of the micelle, and  $B$  is a constant known as the second virial coefficient. The sign of  $B$  provides information on the intermolecular interactions between micelles. A negative value indicates a net attraction between the micelles, and a positive value indicates a net repulsive interaction. A value of zero indicates an “ideal” micellar solution in which there are intermicellar interactions.

#### Example 4.9 Determining the Aggregation Number of a Micelle

Consider values of  $K_o/\Delta R_\phi$  recorded as a function of surfactant monomer concentration. How would you determine the aggregation number of the micelle?

Equation 4.50 can be rearranged to give

$$\frac{K_o(c_m - \text{CMC})}{\Delta R_\phi} 10^{-3} = \frac{1}{M} + 2 \times 10^{-3} B(c_m - \text{CMC})$$

Thus, a plot of the left-hand side of this equation versus  $(c_m - \text{CMC})$  will give a straight line with an intercept equal to  $1/M$  and a slope equal to  $2 \times 10^{-3} B$ .

### 4.10.2 Dynamic Light Scattering

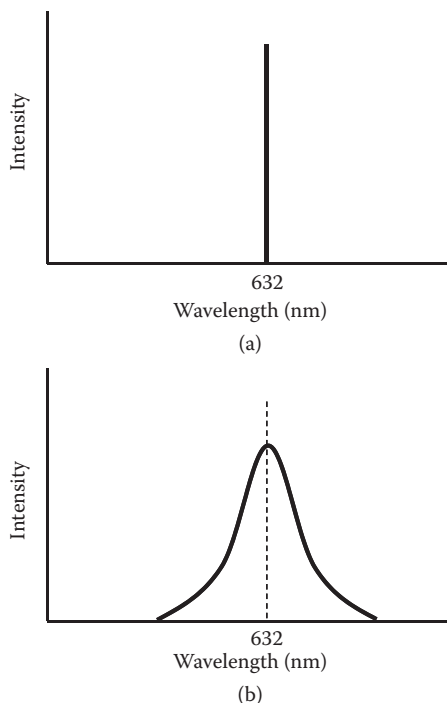
Dynamic light scattering (DLS), also known as photon correlation spectroscopy, is a method used to determine the size distribution of particles in a solution. This is a powerful technique that can accurately report the sizes of particles present across several orders of magnitude. Furthermore, measurements are typically easy to make; this makes DLS ideal for many applications from measuring the size of micelles and other nanoparticles as a function of concentration to assaying a protein solution for the presence of aggregates.

DLS works by shining a laser at a sample solution and observing the scattered light. If the particles in the solution are small compared to the wavelength of the light (if the wavelength is more than 10 or so particle diameters) then this phenomenon will proceed primarily through Rayleigh scattering. In this process, particles absorb a photon and then reemit a photon shortly thereafter. However, the emitted photon is sent in a random direction. Thus, even if the entire incident light comes from one direction, the emitted light is radiated in all directions (scattered). This is an elastic process, meaning that if light of a particular wavelength is absorbed, the scattered light will have exactly the same wavelength. With this in mind, if one conducted an experiment shining a 632-nm HeNe laser at a solution of interest, one might naively expect the scattered light to look like that shown in Figure 4.50(a). This figure depicts all scattered light being detected at exactly 632 nm.

In actuality, however, scattered light is observed in a distribution centered on the expected wavelength [Figure 4.50(b)]. The reason for this is that the particles in solution are undergoing Brownian motion. This means that at any given moment, some particles are traveling toward the detector and have their emitted light blue-shifted relative to the incident light. Similarly, some particles are moving away from the detector and have their emitted light red-shifted. On average, however, the particles are not moving relative to the detector and thus the distribution's center lies on the wavelength of the incident light.

The observed wavelength of scattered light is not the only thing affected by Brownian motion. The total observed intensity also fluctuates in time. This is because Brownian motion causes the distance between any two adjacent atoms to vary in time. The light emitted from these atoms experiences interference; the distance between the atoms determines whether this interference is constructive or destructive. Thus, as this distance varies for a certain pair of atoms, the net light emitted

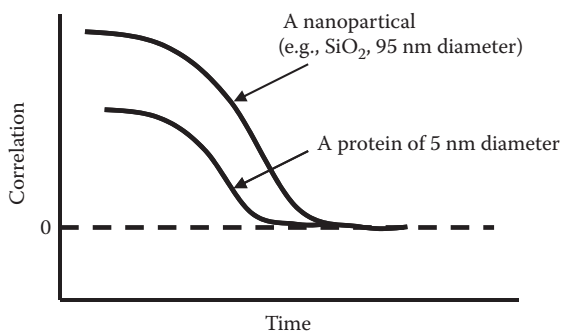




**Figure 4.50** (a) In a simplistic model all scattered light is detected at one wavelength (the wavelength of the incident light). (b) The actual observed scattered light; the wavelengths are distributed around the expected wavelength due to Doppler shifting of the light from Brownian motion of the particles.

from them will fluctuate in intensity. Since there are so many particles in solution, the net intensity of scattered light approximately evens out and is never very far from the average intensity. However, it does not perfectly average out, and the net intensity of signal in solution can be seen to fluctuate.

Critically, the intensity at any given point in time depends on the location of all of the particles in solution. Thus, if one takes a measurement at time zero and another measurement very quickly after that, the intensity of the second measurement will be very similar to that of the first measurement so long as the particles have not had sufficient time to diffuse far from their original positions. If, on the other hand, the particles have had enough time to move significantly, the intensity will be essentially random. This concept can be expressed as a correlation. That is, if the second measurement is likely to be more similar than average to the first measurement, that is a positive correlation. If, however, the second measurement is



**Figure 4.51** A plot of correlation versus time delay between subsequent measurements for two different particles. As expected, the larger particles stay correlated for a longer period of time, indicating that they diffuse more slowly.

likely to be less similar than average to the first measurement, that is called negative correlation. A correlation of zero means there is no relation between the measurements and the first intensity gives the experimenter no information about the second intensity. In a DLS experiment, as the time delay between the two intensity measurements increases from zero to infinity, the correlation decreases from unity (perfect correlation) to zero (no correlation).

We can take this data and plot correlation versus time delay for solutions of different particles (Figure 4.51). This graph gives a time scale for how fast a given particle diffuses. The faster it diffuses, the more quickly it will reach zero correlation as the particles need less time to move to different positions from where they started. In fact, these correlation graphs can be used to calculate the rate of diffusion,  $D$ , which can be employed in further calculations.

The aforementioned rate at which a particle diffuses can be entered into Equation 4.51 to calculate the particle's hydrodynamic radius. In this equation  $d_H$  is the hydrodynamic diameter,  $k_B$  is the Boltzmann constant,  $T$  is temperature,  $\eta$  is the solution viscosity, and  $D$  represents the rate of diffusion (called the translational diffusion coefficient).

$$d_H = \frac{k_B T}{3\pi\eta D} \quad (4.51)$$

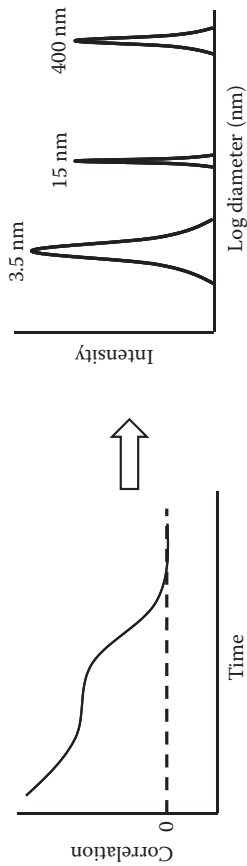
As one might expect, the larger the particle, the slower it diffuses, as shown in the Figure 4.51; this is why the diffusion rate and diameter have an inverse relationship. The number of particles observed at each diameter can be plotted versus

diameter to obtain a graph such as that in Figure 4.52. This figure demonstrates how a complex correlation curve with two distinct exponential regions can be transformed into an easy-to-interpret graph with two peaks at the corresponding particle sizes.

There are several things that can affect diffusion speed besides particle size that need to be controlled for (or at least considered). The first is the size of the hydration shell around each particle. Since it is the hydrodynamic radius of each particle that is measured, if two particles have equal size, but one is more strongly coupled to the surrounding media, that particle will diffuse more slowly and register as larger. Second, the ionic strength of a solution affects the coupling of the media to the particles, so ionic strength must typically be set at a standard level. Third, the surface morphology of a particle can affect diffusion speed. If a particle has long, comb-like appendages on its surface, they will slow its diffusion and make it appear larger than a smooth sphere of equal size. Finally, changing particle geometry can affect diffusion speed. If a particle shifts from a more compact spherical formation to a more extended conformation, it will be slowed down and again appear larger despite not changing mass at all. This last effect is of particular interest in protein analysis because it can sometimes detect protein morphology shifts between different forms.

One important limitation to DLS is that the signal intensity received from a given particle is proportional to the sixth power of the particle's diameter. Thus, if one has a solution containing a 1:1 mixture by molar concentration of 10-nm particles and 100-nm particles, the peak area corresponding to the larger particles will be a million-fold larger than that for the smaller particles despite the solution containing an equal number of each. This problem only worsens as the size difference increases and serves to limit the range of particles that can be accurately analyzed in a single solution.

Dynamic light scattering is an excellent method for determining the sizes of particles present in a solution. It is able to accurately size particles from the nanometer scale to above the micron scale so long as any one particular solution does not contain widely varying particle sizes. This technique has found utility in applications ranging from colloidal science to proteomics, and has become an integral part of many nanoscience laboratories. Few other methods of particle sizing present such an attractive combination of ease of use and wide region of sensitivity.



**Figure 4.52** A plot of signal intensity versus particle size along with the correlation plot from which such a graph is derived.

## REFERENCES AND RECOMMENDED READING

- Evans, D. F. and Wennerström, H. 1999. *The Colloidal Domain*, 2nd ed. New York: Wiley-VCH. Chapter 2 provides a light read on AFM and STM.
- Tompkins, H. G. 2006. *A User's Guide to Ellipsometry*. Dover Publications, Mineola, NY. This is a graduate-level book that covers the theory of ellipsometry in detail and provides many interesting applications of the technique.
- Vickerman, J. C. 2003. *Surface Analysis—The Principal Techniques*. John Wiley & Sons, Chichester, West Sussex, UK. This book provides an excellent coverage of surface science techniques, many of which are not covered in this chapter. The book is recommended for students who wish to gain a better understanding of vibrational spectroscopy at surfaces and scanning probe methods.
- Boyd, R. W. 1992. *Nonlinear Optics*. Academic Press, San Diego, CA. This book is recommended for the advanced student interested in nonlinear methods such as second-harmonic generation and sum-frequency generation. The book provides an excellent fundamental treatment of nonlinear optics.
- Bonnel, D. 2000. *Scanning Probe Microscopy and Spectroscopy: Theory, Techniques, and Applications*, 2nd ed. Wiley-VCH, New York. This book focuses mainly on STM and is only recommended for the student who is seriously interested in this method.
- Berne, B. J. 2000. *Dynamic Light Scattering: With Applications to Chemistry, Biology, and Physics*. Dover Publications, Mineola, NY. This book explains DLS and how Maxwell's equations lead to the intensity of the scattered radiation. This is an advanced graduate-level textbook.

## END OF CHAPTER QUESTIONS

1. Equation 4.1 can be used to determine the surface tension of a liquid by measuring how far the liquid travels up a narrow capillary tube. How far would water rise up a narrow capillary of radius 1 mm? Assume that water makes a zero contact angle against the surface of the tube.
2. A paper plate of width 10 mm and thickness 1 mm is withdrawn from a surfactant solution. A force of 0.77

mN is measured just before the plate detaches from the surface of the solution. Estimate the surface tension of the surfactant solution.

3. In a QCM experiment, the following values of frequency (at the third overtone,  $n = 3$ ) were measured in real time as a nanofilm of mammalian DNA was adsorbed on a quartz crystal with Sauerbrey constant  $C = 17.7 \text{ ng Hz}^{-1} \text{ cm}^{-2}$ .

Time (min)	0	5	10	15	20	25	30	35	40	45	50	55	60	65
$\Delta F$ (Hz)	0	0	-2	-5	-7	-15	-25	-42	-55	-60	-63	-64	-65	-65

Use the Saurbrey equation to generate a plot of time versus adsorbed mass. Assume the adsorption follows first-order kinetics, i.e., [Adsorbed DNA mass] =  $(1 - Ae^{-kt})$  where  $A$  is a constant and  $k$  is the first-order rate constant. Use this equation to obtain a value for  $k$ .

4. Consider solutions of bovine serum albumin (BSA), a common blood protein, made at pH 3, 5, and 7 flowed over a silicon waveguide in a DPI experiment. pH 5 is close to the isoelectric point of BSA (or the pH at which it has no net charge), where it is expected that each BSA protein experiences little electrostatic repulsion with neighboring proteins at pH 5.
- (a) Predict at what pH you would expect the maximum adsorption of BSA. Explain your answer.
- (b) In bulk solution, BSA adopts an ellipsoid shape with dimensions of  $140 \text{ \AA} \times 40 \text{ \AA}$  and it has a molecular weight of 66.43 kDa. Using this information in conjunction with the DPI mass values, how would you calculate the area occupied by each protein molecule on the waveguide surface at each pH?
- (c) If we were to assume that the BSA proteins adsorbed parallel to the waveguide surface to form a saturated monolayer and that they were not distorted upon adsorption, then we could use the bulk phase dimensions of BSA to calculate the nanofilm thickness (it's  $\sim 40 \text{ \AA}$  and each molecule occupies  $\sim 5600 \text{ \AA}^2$  on the surface at pH 5). The nanofilm layer thickness values are much less for the BSA proteins at pH 3 and 7, and the area-per-molecule values are much greater than predicted. Thus, it can be inferred that the BSA

molecules at pH 3 and 7 are spread much more thinly on the waveguide surface and that they assume a much more distorted form. The relative rates of adsorption of BSA at different pH to the silicon surface are  $\text{pH } 3 > \text{pH } 5 > \text{pH } 7$ . What forces are responsible for the adsorption of BSA to the waveguide surface at each pH? What explanation can be offered for the observed order of the relative rates of adsorption of BSA at differing pH?

5. The following table shows the refractive index of poly-ethylenimine (PEI) as a function of concentration. The refractive index of a film of PEI on the surface of a DPI waveguide is 1.55. Assuming this film is under a pure water solution, estimate the density of the film. In a separate experiment, ellipsometry determined this film to have a thickness of 1 nm. What is the mass of the film?

Concentration (mM)	1.0	5.0	10	15	20
Refractive Index	1.44	1.45	1.46	1.47	1.48

6. Common IRE materials used in ATR-FTIR are zinc selenide (ZnSe) and germanium (Ge). Using tabulated values in the scientific literature, we can find the index of refraction values for ZnSe and Ge at a variety of incident light wavelengths. Using other known experimental parameters allows us to calculate the penetration depth of the evanescent wave into the sample being studied. For example, at IR frequencies of  $\sim 1700\text{ cm}^{-1}$  (the typical stretching frequency of a C=O bond), ZnSe has an index of refraction of  $\sim 2.35$ , whereas the RI of Ge is  $\sim 4.0$ . Assuming that the angle of incidence used in an IRE of an ATR-FTIR is  $45^\circ$ , compare the evanescent wave penetration depths in an aqueous sample (RI  $\sim 1.5$ ) for an IRE made of either ZnSe or Ge when IR radiation of  $1700\text{ cm}^{-1}$  is used.

Hint: See the Appendix for a useful equation to calculate the penetration depths.

7. Consider a nanofilm comprised of discrete layers of a chromophore, such as a dye molecule. Ellipsometry can measure the thickness of such a film. If the molar absorptivity of the molecule is known, show how Equation 4.23 can be used to determine the mass of the film.

8. The Sauerbrey equation underestimates the mass adsorbed to the crystal surface in liquids because it was originally developed for oscillation in air and does not account for the decrease in resonant frequency due to the change in viscosity. It is possible to amend the Sauerbrey equation for an accurate measurement of mass in liquid by correcting for these viscosity effects according to the following equation:

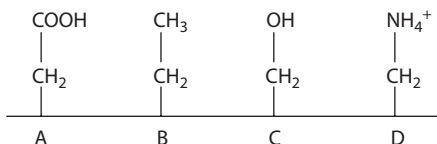
$$\Delta f = -f_o^{3/2} \sqrt{\left( \frac{\eta_l \rho_l}{\pi \rho_q \mu_q} \right)}$$

where  $f_o$  is the resonant frequency.  $\rho_l$  is the density of the liquid.  $\eta_l$  is the viscosity of the liquid.  $\rho_q$  is the density of quartz (2.648 g/cm<sup>3</sup>).  $\mu_q$  is the shear modulus of quartz for AT-cut crystal ( $2.947 \times 10^{11}$  g/cm s<sup>2</sup>). In which liquid, water or deuterium oxide, is the detection limit of a typical QCM-D instrument at its fundamental resonant frequency more sensitive? What is the ratio of their sensitivities? (The interested student should see the following reference for more details: K. K. Kanazawa and J. G. Gordon, II, "The oscillation frequency of a quartz resonator in contact with a liquid," *Anal. Chimica Acta*, vol. 175, pp. 99–105, 1985).

9. (a) The text mentions that many techniques cannot by themselves provide absolute measurements of properties such as refractive index, thickness of a film, or density. Choose two non-spectroscopic techniques discussed in the chapter and describe how their combined use provides more information than either would alone. Outline a set of experiments that you would undertake to learn as much as possible about a particular film-forming substance such as SDS (or provide a compound).
- (b) Choose a spectroscopic method, such as IR, and describe how this technique could add to your understanding of the material and its adsorption properties.
10. Describe how you would use information given about molecular fluorescence spectroscopy to calculate  $\Phi_f$  for a set of data. (Hint: Use an appropriate equation and integrate it to find  $g_x$  and thus calculate  $\Phi_f$ .)

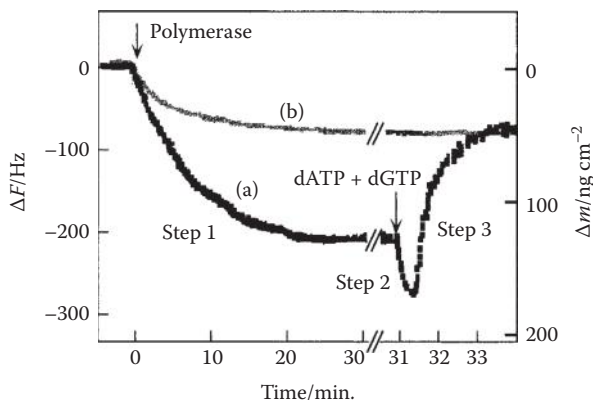


11. Scanning tunneling microscopy is used to image a metal surface with an electronic decay length of  $5 \text{ nm}^{-1}$ . If the tunneling current increases by a factor of 5, then how much has the height changed? Describe what this height refers to in your own words. (Hint: See Example 4.8.)
12. Rank from tallest to shortest the height of an AFM tip that has been functionally coated with  $-\text{COOH}$  as it traverses a surface of the following makeup:



13. I'm using NSOM to detect the incorporation of malaria-causing *Plasmodium* proteins in the membrane of malaria-infected human erythrocytes. I place my sample 300 nm from the NSOM probe, but see that the resulting pictures are not resolved. I then realize that the wavelength of the light source from my laser is 150 nm. What must I do to increase the resolution?
14. What is the concentration fluctuation factor of a surfactant that forms micelles with very large molecular weight? Express your answer in terms of the second virial coefficient.
15. (a) Explain in words why the distribution of Rayleigh scattered light is a broad peak instead of a sharp signal.  
 (b) If a laser emitting light at 700 nm is sent through a dispersion and the Rayleigh scattered light is centered around 640 nm, what can be deduced about the direction of motion of the particles in solution?  
 (c) Does the shift mentioned in part (b) correspond to a red shift or blue shift?
16. (a) If the hydrodynamic radius of a particle doubles, what happens to its diffusion speed?  
 (b) If the radius is halved, the temperature is quadrupled, and the viscosity of solution increases by half, by what factor does the diffusion speed change?

- (c) In a solution of hexane, would one expect to observe multiple signals? Explain why or why not.
17. Explain why the maximum feature resolution in contact lithography is approximately  $\lambda/2$ . (See Section 5.9.)
18. This question concerns SFG spectroscopy. Consider the overlap of two light beams (one visible of wavelength 523 nm and the other IR of wavelength 2.2  $\mu\text{m}$ ) at a solid surface. The angle of incidence of the visible beam is  $30^\circ$  and that of the IR beam is  $40^\circ$ .
- Calculate the angle of the emitted SFG beam.
  - What is the wavelength of the emitted SFG beam?
  - Describe qualitatively the difference between the SFG spectra of SDS, CTAB, hexanol, and decanol in terms of the  $\text{CH}_2$  and  $\text{CH}_3$  bands. Assume that monolayers of these molecules are present between a hydrophobic surface and an aqueous phase.
19. In a study conducted by the Tokyo Institute of Technology, QCM was used to directly monitor the reactions of the Klenow fragment of *Escherichia coli* DNA polymerase I. The reactions take place on DNA oligonucleotides, containing either a  $(\text{TTTTTC})_3$  or  $(\text{TTTTTC})_{10}$  template and the primer necessary for polymerase binding, which have been immobilized on the QCM. This enzyme acts as a catalyst for complementary base pairing of dATP and dGTP monomers from solution into a second strand. The results of two trials with different initial conditions for the  $(\text{TTTTTC})_3$  template are shown below:



In (a) polymerase was added at first, and then excess monomers were added after 30 minutes. In (b) polymerase was added in the presence of excess monomers.

- (a) What is physically occurring at steps 1, 2, and 3 of (a) in terms of the biological components of the reaction? What is physically occurring in (b) that causes a different curve? Why does it make sense that the final changes in frequencies are the same?

Hint: The frequency change in step 1 of (a) is of roughly equivalent magnitude to the frequency change in step 3 of (a), and the frequency change at step 2 of (a) is roughly equivalent to the total frequency change in (b).

- (b) What can be said about the relative magnitudes of the rate constants for each of the three steps in the DNA polymerase reaction mechanism based off the data above?
- (c) What relative changes in frequency, coupled mass, and rate constants would be expected if using the  $(\text{TTTTTC})_{10}$  template instead of the  $(\text{TTTTTC})_3$  template?

## Five

---

# Types and Uses of Some Nanomaterials

---

### CHAPTER OVERVIEW

This final chapter is written more like a scientific review of nanomaterials, with the goal of helping the student to transition from the fundamental textbook-style material to the primary scientific literature. The previous chapters should have given the student the necessary scientific background and terminology to understand the material presented in this chapter. This chapter can be thought of as two separate sections: tuning optical and energetic properties of materials using nanomaterials (Sections 5.1 to 5.4) and functionalizing surfaces using nanomaterials (Sections 5.5 to 5.9), brought together as they both define useful ways to employ nanotechnology. There are many other potential applications of nanomaterials, but herein an overview is given of how to functionally use various technologies, each of which is constantly evolving the state of nanoscience.

### 5.1 SUPRAMOLECULAR MACHINES

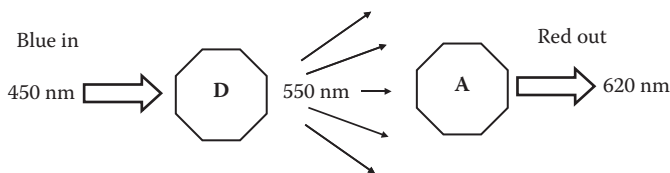
In every self-assembly process highly selective interactions occur that control the transition from molecular scale aggregates to nanoscale entities. As the assembly and characterization of molecular-level components has progressed, the potential for nanoscale systems to be functionalized to perform simple tasks is beginning to be realized. Supramolecular chemistry is concerned with the study of the basic features of these interactions and with their implementation in specially designed non-natural systems. The broad category of all such nanoassemblies has been dubbed supramolecular machines. These machines are comprised of organized systems of cooperating molecules, and are deliberately and specifically engineered with the ability to function as sensors, processors, and custom catalysts. One area of supramolecular fabrication

that is generating considerable interest is the specific programming of energy transfer through an assembly, which is accomplished by a system of interacting dye molecules. This section addresses several systems of interacting dye molecules in nanofilm assemblies and discusses several applications where the careful orientation of donor and acceptor fluorescence partners is exploited for specific film functions. This section was inspired by the excellent coverage of this material in Kuhn and Försterling's textbook, *Principles of Physical Chemistry: Understanding Molecules, Molecular Assemblies, Supramolecular Machines*. Students are encouraged to read Chapters 22 and 23 of this book.

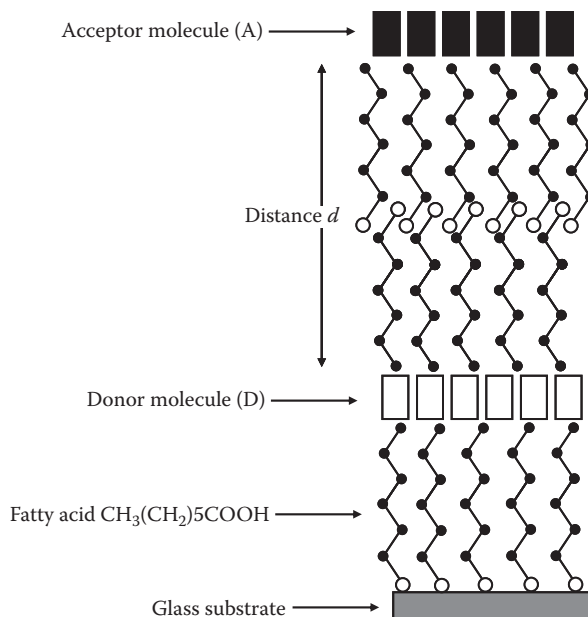
### 5.1.1 Model Dye System

In order for an assembly of molecules to become a useful system for performing work, it is necessary to fabricate an effective means of energy transfer to and/or across the machine. Photon energy is a highly favorable way to deliver such an energy source for a variety of reasons, including the ability to tune the wavelength (color) of the source, minimal invasiveness, and the ability to selectively excite various sites of an assembly. Processes that harness light energy for later use as chemical energy are abundant in nature, and thus make the understanding and fabrication of such devices an important task.

The most basic method by which photon energy can be harnessed by molecular systems is through the exploitation of the specific absorption-fluorescence characteristics of dye molecules. Let us consider a model system of two interacting dye molecules: donor (D) and acceptor (A) dyes, as shown in Figure 5.1. Donor D can be irradiated with light of wavelength 450 nm (blue light), resulting in excitation and fluorescence at 550 nm (green light). A different dye, acceptor A, behaves similarly upon exposure to 550-nm wavelength light by absorbing it and fluorescing red light at 620 nm. We can exploit the overlap of D's fluorescence and A's absorbance in



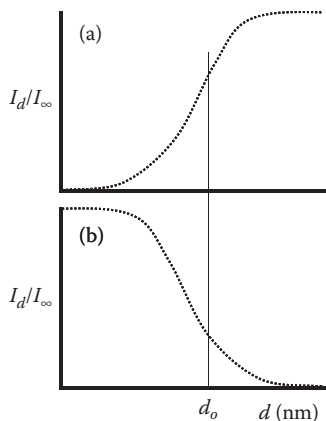
**Figure 5.1** Excitation of donor molecule (D) by blue light (450 nm) and its subsequent fluorescence of green light (550 nm). The emitted green light is absorbed by acceptor molecule (A) which finally emits red light (620 nm).



**Figure 5.2** A layer-by-layer assembly in which a monolayer of donor molecules is separated from a monolayer of acceptor molecules by “inert” fatty acid multilayers. The distance  $d$  is determined by the number of discrete fatty acid monolayers between the A and D monolayer.

order to fabricate a simple molecular machine. Specifically, we can monitor the intensity  $I$  of fluorescence from dye monolayers on fatty acid precursors and spacers to determine their distance of separation,  $d$ . These films can be assembled easily via the hydrophobic/hydrophilic interactions in a Langmuir-Blodgett (discussed later) or electrostatic self-assembly (ESA) deposition process. As shown in Figure 5.2, the hydrophilic fatty acid headgroups associate with a glass substrate, exposing the molecule's hydrophobic tail. Subsequent depositions of dye molecules D, a certain number of fatty acid spacers, and molecules of A result in a simple molecular machine. Using photoquenching equations derived by Drexhege et al. (1963) and later applied to this system by Mobius (1969), the intensity  $I$  of the green (D) and red fluorescence (A) can be calculated at any separation distance  $d$  and verified by monitoring  $I$  in spectroscopic experiments.

Figure 5.3(a) expresses the quantum yield as a ratio of green fluorescence intensity from D when A is located at a distance  $d$  ( $I_{d,D}$ ) to the intensity of green fluorescence when  $d$  is infinitely large or A is absent ( $I_{\infty,D}$ ). This value is a function of



**Figure 5.3** Quantum yield profiles illustrating energy transfer from D to A. (a) Green fluorescence yield from D as a function of distance  $d$ . (b) Red fluorescence yield from A as a function of distance  $d$ . As  $d$  increases, energy transfer to A is reduced resulting in more fluorescence from D and less from A.

the separation distance of the two dyes, where  $d_0$  is the distance where half of the fluorescence from D is quenched by A. Conversely, Figure 5.3(b) expresses the ratio of red fluorescence intensity from A when precursor D is located at a distance  $d$  ( $I_{d,A}$ ) to the intensity of red fluorescence when  $d$  is infinitely large ( $I_{\infty,A}$ ) as a function of  $d$ . On the left,  $I_d/I_\infty$  for donor fluorescence is plotted versus  $d$ . As expected, green fluorescence *increases* with  $d$  since more green fluorescence is transmitted as acceptor dye A becomes less accessible by  $d$  increasing. In Figure 5.3(b),  $I_d/I_\infty$  for acceptor fluorescence is plotted against  $d$ . As expected, red fluorescence *decreases* with  $d$  since A is excited with less efficiency as it becomes more distant from the excitons (the excited regions of the assembly) released from D, which are needed to create fluorescence in A.

These results themselves do not furnish us with much information beyond a confirmation that our theoretical model for energy transfer between the dyes is valid. However, Lehn (in *Supramolecular Chemistry*, VCH Weinheim, 1995) presents a compelling application for such behavior in dye assemblies. Consider two antibodies that both bind specifically and at a different site to an antigen. By coupling dyes A and D to the two different antibodies, the dyes become sufficiently close to allow energy transfer. By introducing an excess of the modified dyes, spectroscopic analysis could be used to perform a simple immunoassay. After accounting for background fluorescence from modified antibodies unbound to the antigen interacting

with one another directly, fluorescence from acceptor A could be monitored and would indicate the specific concentration of the antigen in question.

### 5.1.2 Photorelaxation

The functionality of the aforementioned dye assemblies hinges on the differences in absorbance and emission spectra, or Stokes shift, of the two entities (not to be confused with Stokes shifts in Raman spectra). As a molecule is subjected to energy, in this case energy in the form of photons, it enters an excited, quantized electronic state. This excited state cannot be sustained—in fact, exciton lifetime is approximately 10 nanoseconds—and the molecule quickly reverts to its ground state. Upon returning to ground electronic state, the molecule is left with excess energy from the initial excitation, and must thus emit that energy in several forms. In dyes, energy is dissipated via the emission of a photon. Visible color is emitted from photorelaxing dyes in very specific cases when the energy band gap, or energy difference between ground and excited states, lies between the energies available to photons with a visible wavelength. Since the band gap and emission of a photon from an excited molecule can be quantified by

$$E_{\text{BandGap}} = E_{\text{Photon}} = \frac{hc}{\lambda} \quad (5.1)$$

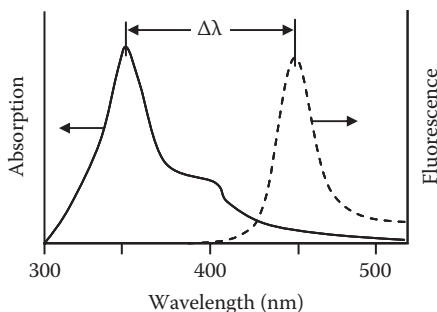
where  $h = 6.62606896 \times 10^{-34}$  J s and  $c = 2.998 \times 10^8$  m/s, we can determine that the energy band gap must exceed  $2.64866 \times 10^{-28}$  J and be less than  $5.22762 \times 10^{-28}$  J (or  $1.65316 \times 10^{19} - 3.26282 \times 10^{19}$  eV) in order to allow for the dye to emit a quantum of visible light. The difference between the excitation photon wavelength necessary to induce an electronically excited state and the wavelength of the subsequently emitted photon is known as the Stokes shift. The Stokes shift is most often positive; in other words, the energy of the exciting photon is partially diminished via other non-radiation energy loss processes (see Figure 5.4) during excitation, and thus the emitted photon is of lower energy (higher wavelength).

#### Example 5.1 Stokes Shift

Estimate the Stokes shift in kJ/mol for the spectra shown in Figure 5.4.

*Solution*  $\Delta\lambda$  is around 100 nm. According to Equation 5.1,





**Figure 5.4** A typical excitation (absorption) and emission (fluorescence) spectrum of a dye molecule. The difference  $\Delta\lambda$  is a measure of the Stokes shift.

$$E = \frac{hc}{\lambda} = \frac{(6.62 \times 10^{-34} \text{ Js})(2.998 \times 10^8 \text{ m/s})}{100 \times 10^{-9} \text{ m}} \approx 2 \times 10^{-18} \text{ J}$$

For one mole we multiply  $E$  by the Avogadro number ( $6.022 \times 10^{23} \text{ mol}^{-1}$ ) and then divide by 1000. This gives 1200 kJ/mol.

In certain materials, however, anti-Stokes shifts occur and are interesting phenomena. At first glance, an anti-Stokes shift seems to break fundamental laws of nature: a lower energy photon is transformed into a higher energy one. Where is this energy coming from? The answer to this question lies in the molecule's ability to remove heat from itself. In materials exhibiting anti-Stokes shift behavior, photon energy is absorbed; however, before reverting to ground state, the exciton also absorbs a quantized mode of vibration. This mode is modeled as a particle, called a phonon. This extra energy absorption creates a higher energy photon emission; however, loss of this vibration energy creates a loss of heat in the molecule. In this way, anti-Stokes shift materials cool as they photorelax.

### 5.1.3 Formation and Properties of the Exciton

When we examine dye molecules 1 and 2 in interaction with one another with 2 deposited on top of 1, we find that the absorption spectrum is different from the sum of their individual spectra. Specifically, dye 1, absorbing at lower wavelengths, is shifted to even lower wavelengths (higher energy), and dye 2, absorbing light at higher wavelengths, is shifted to higher wavelengths (lower energy). In order to explain the cause of this difference, we can approximate the behavior of the dye molecules by oscillating charges exposed to the oscillating

forces of the incident light. When considering our system as a whole, we must account for the interactions among the electrons to be excited. As incident light reaches the system, each  $\pi$ -electron system is either in phase or out of phase with its neighbors. In-phase oscillation shows a strong absorption band because both oscillating charges are accelerated by the electric field of the excitation. In out-of-phase oscillation, a weak absorption band is realized because one charge is accelerated while its neighbor feels an opposing electric field and is decelerated.

However, we must consider our aggregate system of dyes D and A. In an aggregate layer, the oscillator that we use to represent a dye molecule will be strongly coupled to all the other dye molecules (oscillators) in the layer due to the tight packing in the system. Namely, the  $\pi$ -electrons of each molecule to be excited are in close contact with all the other electrons, and their oscillations are in phase. This means that as incident light is exposed to the dye interface, an in-phase oscillation among all the layers' oscillators occurs and an excited domain forms, having absorbed the energy from the incident photons. We refer to this excited domain as the exciton.

We can quantify the approximate lifetime of the exciton in our model system. It has been shown that the fluorescence lifetime of a monomer  $\tau_0$  in terms of our oscillator model is

$$\tau_0 = \frac{3m\epsilon_0 c_0^3}{2Q^2 \pi v_0^2 n} \quad (5.2)$$

where  $m$ ,  $Q$ ,  $v_0$  is the mass, charge, and frequency of the oscillator,  $\epsilon_0$  is the permittivity of the vacuum,  $c_0$  is the speed of light in vacuum, and  $n$  is the refractive index. In the aggregate,  $N$  oscillators are oscillating in phase, contributing to the excited domain. Therefore, we should replace  $m$  by  $Nm$  and  $Q$  by  $NQ$ , obtaining

$$\tau_{agg} = \frac{3(Nm)\epsilon_0 c_0^3}{2(NQ)^2 \pi v_0^2 n} = \frac{1}{N} \tau_0 \quad (5.3)$$

We can exploit the fact that the excited domain must exist through a balance between attractive forces of the in-phase oscillating dipoles (excited portion) and the thermal motion tending to disrupt such a domain in order to find the number  $N$  of molecules contributing to the excited domain. It turns out that the binding energy ( $-\Delta E/N$ ) in a pair of oscillators decreases the energy of the excited state, making it

more favorable. Thermal energy ( $kT$ ) works against the former quantity and makes it less favorable, thereby pushing the oscillator out of phase. Therefore,

$$\frac{-\Delta E}{N} = kT \quad (5.4)$$

It has been found that  $\Delta E = -0.24$  eV. Restating Equation 5.4,

$$N = -\frac{\Delta E}{kT} = \frac{0.24 \text{ eV}}{k} \cdot \frac{1}{T} \quad (5.5)$$

and for room temperature,

$$N = \frac{3000K}{300K} = 10 \quad (5.6)$$

We can now use the number of molecules  $N$  contributing to the excited domain, which we solved for in Equation 5.6, in order to approximate the fluorescence lifetime from Equation 5.3.

$$\tau_{\text{agg}} = \frac{T}{3000K} \tau_0 \quad (5.7)$$

Solving Equation 5.2 for  $m = m_e$ ,  $Q = e$ ,  $n = 1.5$ , and  $v_0 = 0.75 \times 10^{15} \text{ s}^{-1}$ , we find that  $\tau = 5 \times 10^{-9}$  s. This yields

$$\tau_{\text{agg}} = \frac{300K}{3000K} \times 5 \times 10^{-9} \text{ s} = 5 \times 10^{-10} \quad (5.8)$$

We can also approximate the speed  $v$  of an exciton through a monolayer. In order to do so, we must consider the area  $A$  covered by the exciton of width  $L$  during  $\tau_{\text{agg}}$ :

$$A = L_{\text{exciton}} \cdot v \cdot \tau_{\text{agg}} \quad (5.9)$$

Let us envision a situation in which a donor photon has a 50% probability of reaching an acceptor molecule. Under this condition,  $A$  must be equal to the area in which just one acceptor molecule is present. Thus, we redefine  $A$  as  $Za$  where  $a$  is the area covered by a donor molecule and  $Z$  is the number of donors per acceptors and

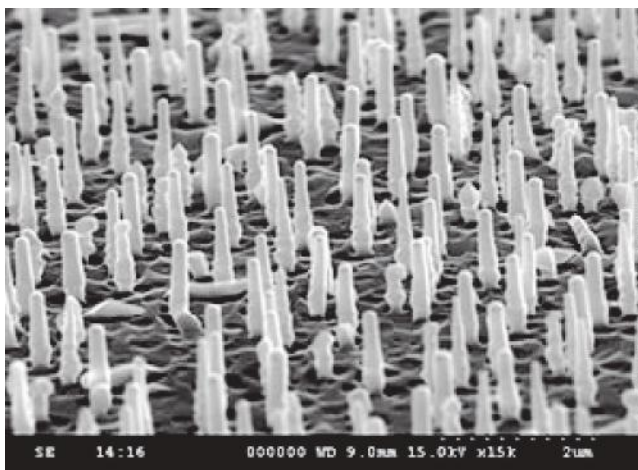
$$v = \frac{A}{L_{\text{exciton}} \cdot \tau_{\text{agg}}} = \frac{Za}{L_{\text{exciton}} \cdot \tau_{\text{agg}}} \quad (5.10)$$

Plugging in typical variable values when the donor is in excess such that  $Z = 10,000$ ,  $a = 1500 \times 400 \text{ pm}^2$ ,  $L_{\text{exciton}} = 5 \text{ nm}$  (at 300 K), and  $\tau_{\text{agg}} = 5 \times 10^{-10} \text{ s}$  we find that  $v = 2.4 \text{ km s}^{-1}$ . For more information see Chapter 23 in Kuhn and Försterling's textbook, *Principles of Physical Chemistry: Understanding Molecules, Molecular Assemblies, Supramolecular Machines*.

Certain experimental observations support a calculation of such a fast-moving exciton. Most importantly, this theoretical number helps explain how the excited domain fluoresces or excites another dye molecule. In this oscillation model, charges in the excited domain change simultaneously and the nearest-neighbor charges are always opposite; thus an attraction of nearest-neighbor charges occurs, leading to a compression of the excited domain. This compression creates a wave that carries the exciton, which does not change shape or lose energy until its destruction through absorption by an acceptor or through fluorescence. This is one possible explanation of how an exciton, or energy, is able to travel through a medium.

## 5.2 NANOWIRES

A nanowire is a nanostructure that has a diameter on the scale of a nanometer and an unrestricted length. An example of a nanowire is shown in Figure 5.5. Nanowires have been synthesized to be as long as one millimeter ( $10^{-3} \text{ m}$ ), but are more commonly made to be about one micrometer ( $10^{-6} \text{ m}$ ). Regardless, the typical nanowire has a width-to-length ratio



**Figure 5.5** A vertical silicon nanowire array on a silicon substrate. (Image provided by Professor Hari Reehal, London South Bank University.)

of 1:1000 or more, allowing the structure to be viewed as one-dimensional (1-D). Combined with the comparatively small size of nanowires, the 1-D structure is responsible for the unique properties that grossly distinguish nanowires from the average (three-dimensional) wire.

The potential applications for nanowires are immense. As technology continues to advance, the trend in industry is to generate smaller but more powerful electronic devices than previously sold on the market. As a result, computer components such as resistors, capacitors, and circuits continue to shrink. To link these components, increasingly small wires are desired. There is also great interest in nanowires from the battery industry, which believes that the increased surface area of condensed nanowires relative to bulk phase metals is the future of anode and cathode design. Industries also see the potential for nanowires in such applications as chemical composites, field emitters, and biomolecular nanosensor leads. Moreover, nanowires' conductance can be controlled by synthesizing the wire from any one of a variety of substrates, allowing for nanowires to have more specialized roles. As a result, nanowires are categorized as (in order of increasing conductance ability) insulators ( $\text{SiO}_2$ ,  $\text{TiO}_2$ ), semiconductors (Si, GaN, InP), or metallics (Au, Cu, Pt, Ni, etc.). With such major applications, research relating to the general properties and synthetic strategies of nanowires is in high demand.

### 5.2.1 Basic Quantum Mechanics of Nanowires

It is well understood that energy is quantized. That is to say, substances are limited to having a discreet amount of energy at any given time. Since the difference in energy between allowed energy states is typically small, large substances can neglect the effect of quantum mechanics. However, quantum mechanics plays an increasing role as devices decrease in size. Nanowires' small sizes result in quantum mechanics having a significant effect on their behavior and properties. This is because quantum mechanics results in substances only being able to absorb or release specific values of energy such that they can maintain their allowed value of total energy. In semiconductors and insulators, the differences in energy levels are often significant and can inhibit electron flow and lower conductance. Metals are different from insulators and semiconductors because for metals the differences in energy states are often negligible. This property of metals results in nearly free electron flow from metal molecule to metal molecule.

Since most applications of nanowires are related to electronics, the quantum effect on nanowires' conductance is of particular interest. Quantum mechanics needs to relate the effect of the discrete energy levels intrinsic of the nanowire molecules to the nanowire's overall conductance. The equation that is used to relate conductance with discrete energy levels is the von Klitzing equation (Klitzing et al., 1980):

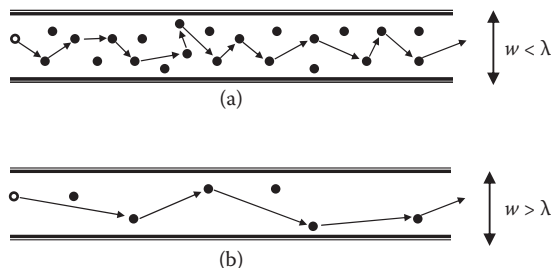
$$R = \frac{h}{ne^2} \quad (5.11)$$

where  $R$  is resistance and is measured in ohms ( $\Omega$ ),  $h$  is Planck's constant,  $e$  is the charge of an electron, and  $n$  is an integer ( $n = 1, 2, 3, \dots$ ). Conductance is simply equal to the inverse of resistance. Therefore, a nanowire has, at  $n = 1$ , a conductance of  $0.0387/\text{k}\Omega$ .

### 5.2.2 Conductivity

As already mentioned, quantum mechanics has great control over conductivity. Yet conductivity varies significantly depending on the composition and the physical properties of the nanowire (width and length). A nanowire's extremely large length-to-width ratio results in high levels of resistance. This resistance is compared to that of a bulk phase wire, composed of the same substances, whose resistance is much lower due to the much smaller length-to-width ratio. The nanowire's narrow width contributes in a number of ways to increasing resistance. For one, decreased width increases the relative effect of the wire's surface molecules. The molecules at the surface of the wire are much less tightly bound than molecules found in the interior of the wire. Also, surface molecules are in contact with far fewer of the wire's molecules than those found in the interior of the wire. The overall consequence is that molecules found at the surface of a wire have greater difficulty propagating electrons to other wire molecules. Therefore, high surface-area-to-volume ratios tend to increase the resistance. The basic concept of width effects on resistance is related to what is known as the mean-free path.

Mean-free path relates to the distance an electron travels between subsequent collisions with other moving particles. Collisions are not favorable for electrons in nanowires since the electron can be deflected in any direction (Figure 5.6) and impact the wire's forward flow of electrons. Collisions occur most frequently when the width of the nanowire is smaller than the mean-free path, resulting in frequent collisions and



**Figure 5.6** A schematic demonstrating the effects of mean-free path. The open circle indicates the electron and the solid black circles represent larger atoms.  $w$  is the width of the wire, and  $\lambda$  is the mean free path of the conduction electron. (a) Diffusive transport occurs when the width of the wire is smaller than the mean-free path of the electron, and (b) ballistic transport occurs when the width of the wire is greater than the mean-free path.

impaired wire conductivity. In the opposing case, a nanowire is said to be a “ballistic transporter” when the mean-free path is less than the width. The result is electrons always moving in a forward direction. Ballistic transporters (typically metal nanowires) have very high conductance values compared to those whose wire widths fail to exceed the mean-free path (Takayanagi et al., 2001).

### 5.2.3 Nanowire Synthesis

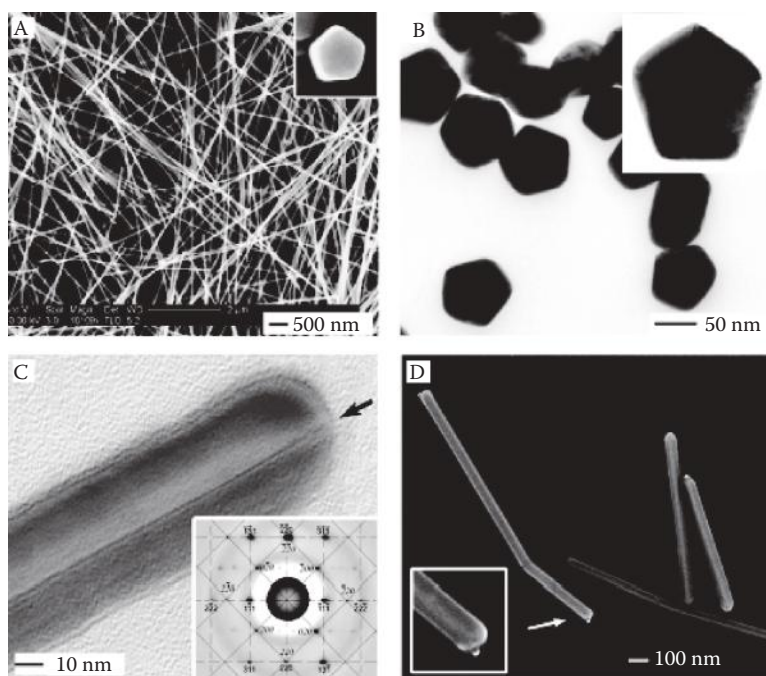
Two general concepts exist for nanowire synthesis, each with their own unique synthetic strategies. The “top-down” approach is one such concept. In this approach, wires are synthesized starting from a bulk substrate and chiseled down until only a field of nanowires remain. Nanolithography and electrophoresis are examples of the top-down techniques employed to make nanowires. Although the theory of the top-down technique is sound, our current resources limit the approach. It is difficult using current instruments to synthesize ideal one-dimensional nanowires. The instruments and techniques are currently limited by the length and width of the nanowires they can synthesize as well as the general accuracy they exhibit when working on a substrate. For these reasons, the top-down approach is presently less commonly employed than its alternative.

The “bottom-up” approach synthesizes nanowires by continuously extending a thread of bound molecules. Various techniques have developed using the bottom-up approach including suspension-based techniques, where a nanowire is elongated in a vacuum using chemical etching or ion-bombardment techniques, and vapor-liquid-solid (VLS) techniques, where a

nanowire is synthesized via a catalyst converting a vapor into a solid below the catalyst, continuously elongating the chain of solid molecules (Crossland et al., 2007). Solution-based bottom-up techniques hold particular promise since they allow scientists to synthesize nanowires in much larger numbers than alternative methods. One particularly intriguing solution-based technique uses copolymers to form templates inside which nanowires can be synthesized (Figure 5.7). Certain copolymers are ideal for template formation because, at the correct temperature and electric field strength, they form evenly spaced hexagonal pores. Once the substrate is added, these pores serve as the locations for nanowire synthesis. Mild reagents can then be used to degrade the template and isolate the nanowires.

### 5.2.4 Summary

Nanowires have great potential. As the field grows, we will begin to see new synthetic techniques capable of producing



**Figure 5.7** Gold nanowires were synthesized using a solution-based bottom-up technique utilizing a copolymer template. Images were taken using a scanning electron microscope. (Reprinted with permission from Chen, Jingy, Benjamin J. Wiley, and Younan Xia. "One-Dimensional Nanostructures of Metals: Large-Scale Synthesis and Some Potential Applications." *Langmuir* 8, 2007: 4120–4129. © 2007 American Chemical Society.)



highly elongated nanowires in large quantities. These techniques will be exceedingly precise, allowing for wire shape and function to be maintained during the entirety of the nanowire synthesis. As the study of nanowires continues to progress, new imaging devices will be forced to develop in parallel to nanowire synthetic techniques. These new devices will allow nanowires to be seen in higher magnification and clarity, allowing for more accurate analysis of experimental results. Studies will continue to be done on synthesized nanowires to help define the nanoparticle's unique properties so it can be better applied to products in the industry including batteries, computers, and chemical composites.

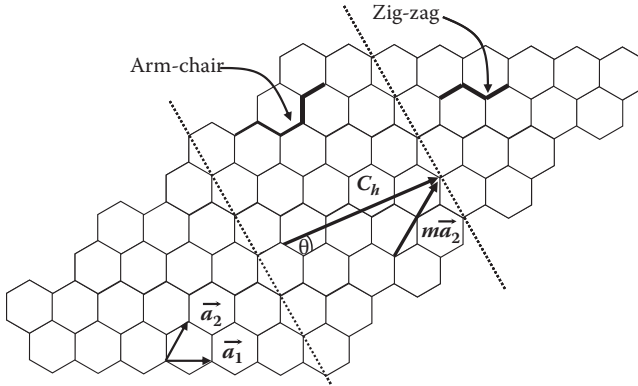
### 5.3 CARBON NANOTUBES

Carbon nanotubes are constructs of carbon that have diameters on the order of a few nanometers and lengths that can be significantly longer. They are extremely strong; the tensile strength of a single-walled carbon nanotube is many times stronger than steel. They also have unique electrical and optical properties and are efficient conductors of heat. These properties have made the study of carbon nanotubes surge in recent years (Dai et al., 2002).

#### 5.3.1 Carbon Nanotube Structure

Carbon nanotubes are a member of the fullerene family of carbon structures; they may contain hexagonal, pentagonal, and sometimes heptagonal rings of pure carbon. It may be helpful to imagine a flat “sheet” of such rings that can be rolled up to make a single-walled nanotube. Regardless of the number of carbons within a ring, each shares a bond with three adjacent carbons. This makes every carbon within the tube  $sp^2$  hybridized. Carbon nanotubes are often capped with a carbon hemisphere that is also composed of  $sp^2$  hybridized carbons in rings. Carbon nanotubes can come in single-walled or multi-walled form. The multi-walled nanotube structure can be thought of as concentric single-walled nanotubes encompassing one another.

Because differing carbon nanotubes may consist of similar  $sp^2$  hybridized carbons, one might assume they all will have similar properties. This is an incorrect assumption. Differences in how the carbon atoms are structurally organized play a significant role in determining the properties of the carbon nanotube as a whole. For example, one way of organizing the carbons can make the nanotube metallic and conducting, whereas another



**Figure 5.8** Schematic diagram showing how a hexagonal sheet of graphite can be rolled to form a carbon nanotube. The properties of the nanotube will depend on the chiral vector and the chiral angle. The unit vectors are shown in the lower left-hand side of the diagram.

organization scheme will make it a semiconductor. Because of this, the atomic structure of a nanotube is often described in terms of its chirality, which can be broken down into a chirality vector,  $C_h$ , and a chiral angle,  $\theta$  (Figure 5.8).

The length of the chirality vector is related to the diameter of the nanotube it describes and can be easily determined if the interatomic spacing of the carbon atoms is known. The interatomic spacing of carbon atoms is often a known parameter because it is the same for all nanotubes with a similar lattice, or unrolled “sheet” of rings. An equation for describing the chiral vector is given in Equation 5.12 (Thostenson et al., 2001).

$$C_h = na_1 + ma_2 \quad (5.12)$$

In this equation  $n$  and  $m$  are the number of steps in the direction of the unit vectors,  $a_1$  and  $a_2$ , in their respective directions across the sheet of unrolled carbon rings. Figure 5.8 is a graphical depiction of this along a graphene sheet, composed of only hexagonally ordered carbon atoms. The chiral angle,  $\theta$ , is a measurement of the resulting twist in the “rolled up” nanotube and is defined as the angle between the chiral vector and the unit vector  $a_1$ . This angle can be determined by relating the number of steps  $n$  and  $m$  as shown in Equation 5.13 (Saito et al., 1998).

$$\cos\theta = \frac{(2n + m)}{[2(n^2 + m^2 + nm)]} \quad (5.13)$$

Nanotubes with chiral vector steps  $n = m$  are known as “arm-chair” nanotubes and have a chiral angle of  $30^\circ$ . Nanotubes with one of the step integers equaling zero are referred to as “zig-zag” nanotubes and have a chiral angle of  $0^\circ$ . Both of these are achiral nanotubes, because their mirror images are identical to the original structure. In all other cases, where  $n \neq m$ , the nanotubes are chiral, with a chiral angle between  $0^\circ$  and  $30^\circ$ .

### 5.3.2 Some Properties of Nanotubes

Interestingly, the relationship of steps,  $n$  and  $m$ , can be used to determine if the nanotube will be conducting or semiconducting. If  $n - m$  is a positive multiple of three, or zero, the nanotube has no band gap, and therefore is a conductor. If  $n - m$  is not a positive multiple of three, or zero, then the nanotube will have a small band gap and therefore is a semiconductor.

The ability of carbon nanotubes to conduct heat is dependent on temperature and is comparable to that of diamond or a graphite sheet. The peak thermal conductivity for a single-walled nanotube is reached at around 100 K and is around  $37,000 \text{ Wm}^{-1}\text{K}^{-1}$ . From there the thermal conductivity drops as the temperature increases to around  $3000 \text{ Wm}^{-1}\text{K}^{-1}$  at a temperature of 400 K.

The precise values for mechanical and elastic properties of carbon nanotubes are currently disputed. They are thought to be the strongest and stiffest materials known to man in terms of tensile strength and elastic modulus. It has been shown that they are many times stronger than steel and that they can return to their original state after being bent. Nanotubes have also been found to be stable at very high temperatures, nearly 3000 K in a vacuum and 700 K in air.

### 5.3.3 Methods for Growing Nanotubes

There are three primary methods for nanotube production: arc discharge, laser ablation, and chemical vapor deposition (Baddour et al., 2005). Arc discharge uses a high-voltage arc to vaporize a carbon source. This frees carbon, which then accumulates into nanotubes in the growth chamber. Following accumulation the nanotubes are collected and purified. Laser ablation is similar to the arc discharge method, but it makes use of a high-energy laser, instead of a current, to vaporize the carbon source. Chemical vapor deposition relies on the decomposition of a carbon-containing gas in the presence of a metal catalyst, such as iron, cobalt, or nickel, which helps facilitate carbon accumulation and nanotube growth. Each method has

many variations with their own strengths and weaknesses that should be considered when choosing which to use.

#### **5.3.3.1 Arc Discharge**

Two carbon rods are set up as an anode and cathode and are placed in a closed growth chamber. The chamber is filled with an inert buffer gas such as helium, argon, nitrogen, or hydrogen and kept at relatively low pressures. Interestingly, the inert gas in the chamber has been found to affect the diameter of grown nanotubes while the pressure in the growth chamber does not. Keider and Waas (2004) found that when mixtures of helium and argon are used to fill the chamber, the molar fraction of argon determines the nanotube diameter. Once the chamber is filled with gas, a direct current is applied until a stable electric arc is formed between the carbon rods. Once the anode reaches a critical temperature, it vaporizes into a gaseous carbon source. The gas in the chamber causes the carbon to deposit. The structures found in these deposits are nanotubes, which need to be purified after collection. This method has been used for the creation of both single-walled and multi-walled nanotubes. In order to produce substantial amounts of single-walled carbon nanotubes a metal catalyst needs to be added to the electrodes.

Unfortunately, once collected, the nanotubes must be manipulated onto substrates for further use, which can be difficult for large quantities of such small structures. Arc discharge also requires a large amount of power and is operated at a high temperature, making it an unfavorable method for the production of large quantities of nanotubes.

#### **5.3.3.2 Laser Ablation**

Laser ablation uses a high-energy laser beam to vaporize a carbon source. The carbon source is sealed off in a quartz tube and heated to temperatures ranging from a few hundred degrees to just over a thousand degrees Celsius. The tube is filled with an inert gas that flows toward a cooled collector at its end. As the laser vaporizes the carbon source, the flow of inert gas carries it down the tube, where it deposits. The nanotubes are then collected from the deposits and purified.

In both methods the vaporized carbon is released as low molecular weight carbon, which coalesces to form larger molecular weight structures. The carbon begins to condense very quickly after vaporization. Using metal catalyst particles doped into the carbon source prevents carbon from closing into a cage structure and decreases the amount of carbon

deposited in non-nanotube form by aggregating with the carbon quickly. This results in more single-walled nanotubes being formed when a metal catalyst is used. Nanotube length can be controlled by changing the time the carbon spends in the high temperature area of the growth chamber, which can be controlled by using heavier or lighter gases and altering the flow rate. Research is undergoing into changing growth conditions, which may make laser ablation a viable option for the large-scale production of nanotubes.

#### **5.3.3.3 Chemical Vapor Deposition**

Chemical vapor deposition relies on the decomposition of a carbon-containing gas, such as methane, ethane, benzene, acetylene, or ethylene, which then aggregates on a catalyst or substrate to form nanotubes. Studies have shown that gases with saturated carbon bonds produce structures with more walls than those grown from unsaturated gases. Therefore, methane and ethane are preferred for single-walled carbon nanotube growth, and ethylene, benzene, and acetylene are preferred for multi-walled carbon nanotube growth. Because high temperatures create more high-quality crystals by reducing defects and increasing crystallization, the source gas must also be one that does not thermally decompose into amorphous carbon and reduce the purity of nanotubes. The amorphous carbon can accumulate on the aggregation sites and inactivate them from forming nanotubes. The aggregation sites are typically metal catalysts on a substrate, but can be a dust or gas if raw material not on a substrate is desired. The size of the catalyst on the substrate determines the diameter of the grown nanotubes.

The nanotubes are similar to those produced by other methods; however, because the carbon is in the gas phase already it does not have such a limited carbon source size as laser ablation or arc discharge methods. There is also more control over the aggregation sites where the carbon nanotubes are grown, meaning that they can be grown as a raw material or on a substrate. Unfortunately, the lower temperatures of chemical vapor deposition result in nanotubes that are not as high quality crystalline as those produced in the laser ablation and arc discharge methods, although techniques have been designed to help overcome this drawback and make this method viable for the large-scale production of high-quality nanotubes.

#### **5.3.4 Catalyst-Induced Growth Mechanism**

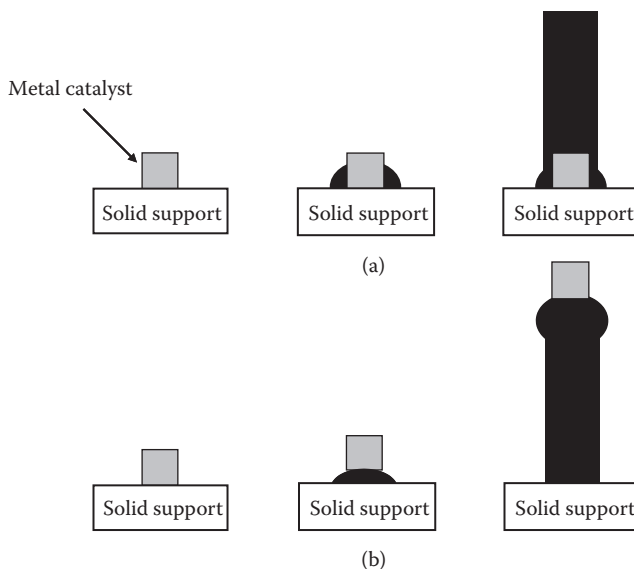
Nanotube growth does not start immediately after carbon gas comes into contact with the metal catalyst. Carbon first

dissolves into the metal catalyst particle. Once the particle is saturated, the carbon atoms assemble into an  $sp^2$  hybridized structure at a less reactive surface. The carbon then moves out of the metal, adding more carbon atoms to the nanotube growing on the less reactive surface at the edges of the particle. This is why the diameter of the metal particle determines the diameter of the nanotube being grown.

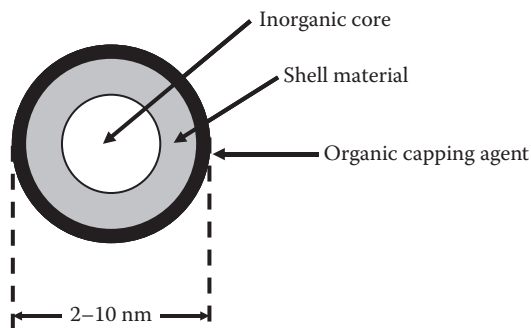
There are two suggested growth mechanisms for nanotube growth from a metal catalyst. The first method is root growth in which the nanotube grows upward from the metal catalyst particle, which stays attached to the substrate. The second suggested growth mechanism is tip growth in which the nanotube grows from the catalyst downward into the substrate. This pushes the catalyst up and detaches it from the substrate (Sinnot et al., 1999). The two mechanisms are illustrated in Figure 5.9.

## 5.4 QUANTUM DOTS

Quantum dots are inorganic semiconductor nanoparticles, typically 2–10 nm in size. The excitons (electron-hole pairs) of quantum dots are confined to three dimensions. Quantum wires are confined to two dimensions, and quantum wells



**Figure 5.9** The two suggested models of carbon nanotube growth through chemical vapor deposition onto a metal catalyst. In both models, the carbon grows around the edges of the metal catalyst.



**Figure 5.10** Structure of a quantum dot. The size and nature of the inorganic core and the shell material can be varied. Quantum dots are often capped with an organic thin film.

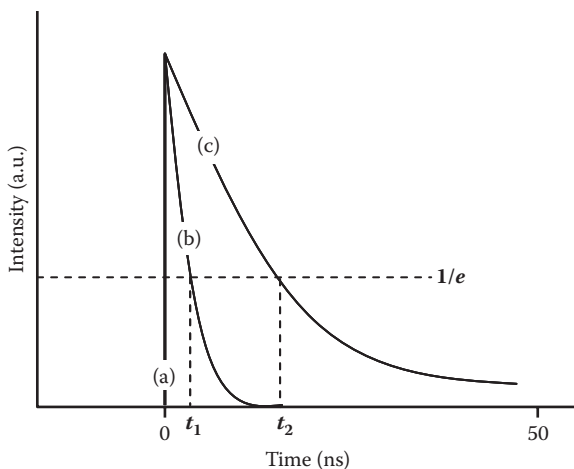
are confined to one dimension. As the confining dimensions decrease toward a limit, the energy gap starts to change with it. This tends to happen at the nanoscale, thus the small size of quantum dots gives them optical and electronic properties in between those of bulk semiconductors and discrete molecules (Brus, 2007). Each dot is a mesoscopic entity with individual properties, differentiating it from a colloidal particle.

Quantum dots typically consist of a core, shell, and final coat, as shown in Figure 5.10. They are characterized by the nature of each of these layers, their size and aspect ratio, their quantum efficiency in optical materials, and their coercivity in magnetic materials. Properties of an ensemble of quantum dots are additionally determined by particle size distribution and differences in morphology within the ensemble.

#### 5.4.1 Optical Properties

Due to quantum confinement, quantum dots have characteristic optical properties that can be used to optimize signal-to-background ratios. High molar extinction coefficients (of the order of  $0.5 - 5 \times 10^6 \text{ M}^{-1} \text{ cm}^{-1}$ ) make quantum dots particularly bright probes when light intensities are severely attenuated by scattering and absorption—their absorption rates are faster than those of organic dyes at the same excitation photon flux. Due to the increased rate of light emission and the large absorption cross-section of quantum dots, they have been found to be 10–20 times brighter than organic dyes.

Additionally, the broad absorption profiles of quantum dots allow simultaneous excitation of multiple wavelengths; thus molecular and cell targets can be tagged with different colors. The fluorescence emission wavelengths can be continuously tuned from 400 to 2000 nm by varying particle size



**Figure 5.11** Comparison of the excited state decay curves (monoexponential models) between quantum dots and common organic dyes. Line (a) indicates excitation. Lines (b) and (c) indicate the decay curves for the organic dye and quantum dot, respectively. A measure of the fluorescence lifetime is indicated by time  $t_1$  (for the dye) and  $t_2$  (for the quantum dot).

and chemical composition. For example, if one compares the emission spectra of rodent skin and that of quantum dots obtained under the same excitation conditions, one finds that the quantum dot signals can be shifted to a spectral region where autofluorescence is reduced in order to improve signal-to-background ratios. Moreover, quantum dots are remarkably resistant to photobleaching and are thus well suited for continuous tracking studies over a long period of time. Signal-to-background ratios can be further improved by time-domain imaging, which is used to separate quantum dot fluorescence from background fluorescence due to the longer excited states of quantum dots (Figure 5.11).

Quantum dots exhibit Stokes shift—the absorption of light at one-wavelength leads to emission at a slightly lower energy, as seen in Figure 5.3. This difference between the excitation and emission peaks can be as large as 300–400 nm, making the quantum dot signal easily recognizable above the background.

## 5.4.2 Synthesis of Quantum Dots

### 5.4.2.1 Precipitative Methods

This method is similar to colloidal synthesis in that it involves supersaturation of a solution followed by crystallization. Like traditional chemical processes, colloidal semiconductor nanocrystals are synthesized from precursor compounds



dissolved in solutions. The precursors are heated to temperatures high enough to convert them to monomers. Once the monomers reach a high enough supersaturation level, the nanocrystal growth starts with a nucleation process. The temperature is a critical factor in determining optimal conditions for crystal growth—it must be high enough to allow for the rearrangement and annealing of atoms during the synthesis process while being low enough to promote crystal growth. Furthermore, the size of the crystal depends on temperature, i.e., it is larger at higher temperatures. Another critical factor is the monomer concentration. At high monomer concentrations, the critical size of the crystal is relatively small, resulting in the growth of nearly all particles and thus facilitating monodispersity. When the monomer concentration is depleted during growth, the critical size becomes larger than the average size present.

The preparation of CdS quantum dots requires the controlled nucleation of CdS in dilute aqueous solutions of cadmium sulfate and ammonium sulfide. The dynamic equilibrium between solvated ions and solid CdS in acetonitrile as a solvent in the presence of a styrene or maleic anhydride copolymer allows the preparation of stable CdS nanoparticles (average size 3.4–4.3 nm). To obtain highly monodispersed nanoparticles, post-preparative separation techniques like size exclusion chromatography or gel electrophoresis are used.

Weller et al. (1985) injected phosphine ( $\text{PH}_3$ ) into solutions containing metal salts to make particles of  $\text{Zn}_3\text{P}_2$  and  $\text{Cd}_3\text{P}_2$ . The particle size can be controlled by varying temperature and phosphine concentration. Even though this method is cheap, it has several disadvantages. It lacks reaction control, which is problematic in large-scale syntheses. Furthermore, important semiconductors like GaAs or InSb are not available for use in this method as they are air and/or moisture sensitive.

#### **5.4.2.2 Reactive Methods in High-Boiling-Point Solvents**

Solutions of dimethylcadmium,  $(\text{CH}_3)_2\text{Cd}$  (in tri-*n*-octylphosphine, TOP), and tri-*n*-octylphosphine selenide (TOPSe) in hot tri-*n*-octylphosphine oxide (TOPO) at 120–300°C produce TOPO-capped nanocrystals of CdSe. Hydrophobic organic molecules like TOPO serve as a reaction medium and also coordinate with unsaturated metal atoms on the QD surface to prevent the formation of bulk semiconductors.

The TOPO method is one of the most widely used methods of quantum dot synthesis as it has several advantages

over other methods. The size of the dots can be controlled by temperature, i.e., they are larger at higher temperatures, and monodispersity can easily be achieved. Additionally, the yields of the reactions are usually high, producing hundreds of milligrams in a single experiment. Moreover, this method is readily adapted to the production of core-shell structures with materials having high quantum efficiencies.

#### **5.4.2.3 Gas-Phase Synthesis of Semiconductor Nanoparticles**

This method involves the atmospheric or low-pressure evaporation of either powders of the pre-formed semiconductor, or co-evaporation of two elemental components (like Zn metal and sulfur). It is not a very favorable method due to the large size distribution of the particles (10–200 nm) and the tendency of particles to aggregate due to the absence of a surface-capping agent.

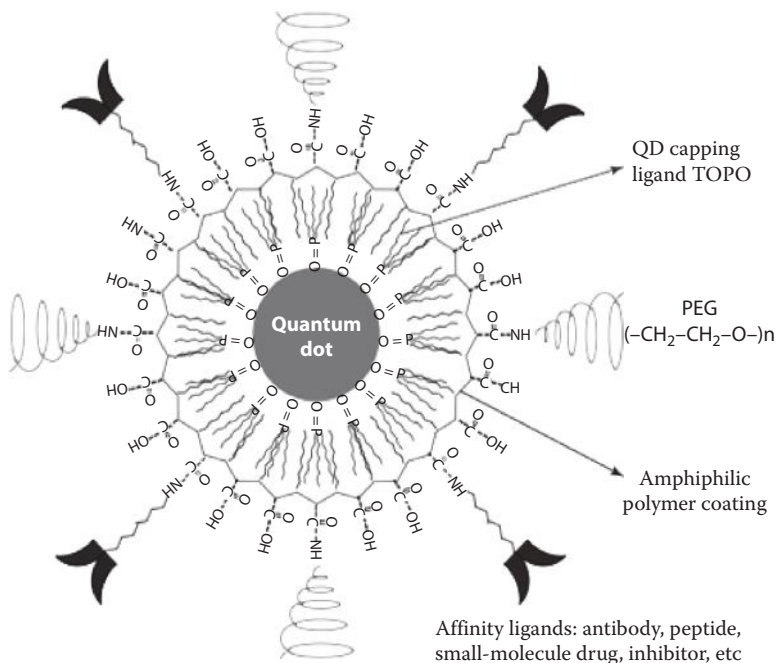
#### **5.4.2.4 Synthesis in a Structured Medium**

A matrix is used to define a reaction space—it provides a mesoscopic reaction chamber in which the crystal can grow only to a certain size. The properties of the system control the properties as well as the size of nanoparticles. Substances typically used as matrices are layered solids, micelles/microemulsions, gels, polymers, and glasses. This technique is often used in biologically related processes. An example of this is the use of empty polypeptide cages found in the iron storage protein ferritin in the synthesis of bio-inorganic nanocomposites of CdS-ferritin.

### **5.4.3 In Vivo Molecular and Cell Imaging**

Due to their optical properties as discussed earlier, quantum dots are better than organic dyes for the purpose of in vivo molecular and cell imaging. Researchers have achieved real-time visualization of single molecule movement in single living cells using quantum dots. Their high electron density allows correlated optical and electron microscopy studies of cellular structures. Quantum dot bioconjugation can be achieved by passive absorption, multivalent chelation, or covalent bond formation as shown in Figures 5.12 and 5.13 (Gao et al., 2005).

Streptavidin-coated quantum dots preferentially bind to biotin with a  $K_D$  of approximately  $10^{-14}$  M, one of the strongest covalent bonds known in nature. Such bioconjugation is widely used for molecular and cellular imaging. An application of this is seen in an experiment conducted to identify  $\beta$ -sheet aggregate-binding ligands using one-bead one-compound screening. The ligands under study were each attached

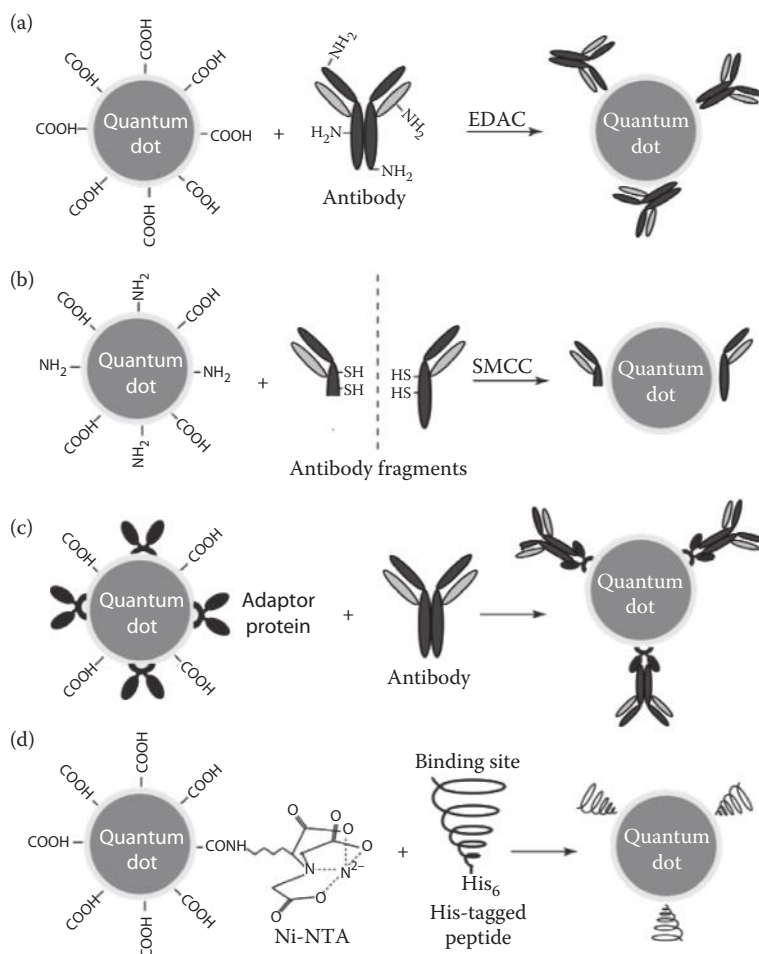


**Figure 5.12** The structure of a multifunctional quantum dot probe. Schematic illustration showing the capping ligand TOPO, an encapsulating copolymer layer, tumor-targeting ligands (such as peptides, antibodies, or small-molecule inhibitors), and polyethylene glycol. (Reprinted from Gao, X.; Yang, L.; Petros, J. A.; Marshall, F. F.; Simons, J. W.; Nie, S. *In vivo* Molecular and Cellular Imaging with Quantum Dots. *Current Opinion in Biotechnology* 2005, 16, 63–72. With permission from Elsevier.)

to a bead. When treated with the biotinylated peptide, only one of these three ligands binds to it. The streptavidin-coated quantum dots preferentially bind to this peptide-ligand combination, and thus the ligand is easily detected due to the fluorescence exhibited by the quantum dots. Analysis by mass spectroscopy reveals the identity of the ligand.

## 5.5 LANGMUIR-BLODGETT FILMS

First demonstrated by Katharine Blodgett in 1934 (Blodgett, 1935), the Langmuir-Blodgett (LB) technique has been by far the oldest and most extensively studied organic nanofilm formation method. Essentially, a monolayer at the air-water interface (the Langmuir film) is mechanically transferred to a solid substrate generating the LB film. Although interest waned until the 1970s, the potential application of LB films in the field of optics and material science has revived LB research.

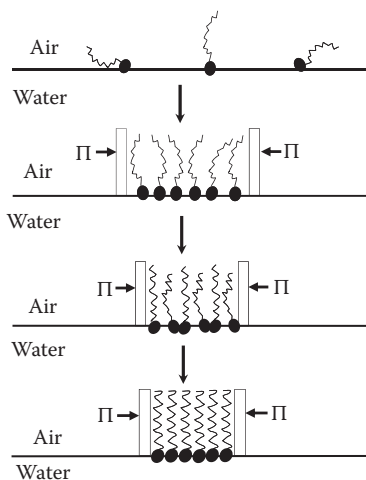


**Figure 5.13** Methods for conjugating QDs to biomolecules. (a) Traditional covalent cross-linking chemistry using EDAC (ethyl-3-dimethyl amino propyl carbodiimide) as a catalyst. (b) Conjugation of antibody fragments to QDs via reduced sulfhydryl-amine coupling. SMCC, succinimidyl-4-N-maleimidomethyl-cyclohexane carboxylate. (c) Conjugation of antibodies to QDs via an adaptor protein. (d) Conjugation of histidine-tagged peptides and proteins to Ni-NTA-modified QDs, with potential control of the attachment site and QD:ligand molar ratios. (Reprinted from Gao, X.; Yang, L.; Petros, J. A.; Marshall, F. F.; Simons, J. W.; Nie, S. *In vivo* Molecular and Cellular Imaging with Quantum Dots. *Current Opinion in Biotechnology* 2005, 16, 63–72. With permission from Elsevier.)

This was fueled by the results that showed an LB film comprised of a dye to be the first multilayer film to produce interesting nonlinear optical (NLO) effects. Nowadays, research toward the synthesis of tailor-made molecules with intrinsic nonlinear susceptibilities and their incorporation into LB films is flourishing.

### 5.5.1 Langmuir Films

The first step of LB multilayer assembly is to form a Langmuir film by dissolving surface-active organic molecules, such as surfactants or amphiphiles, in a nonpolar, volatile solvent and then spreading the solution onto a polar liquid surface, usually water. Once the volatile solvent evaporates, the remaining amphiphilic molecules are oriented at the air-water interface so that the hydrophilic headgroups are buried in the bulk aqueous phase while the hydrophobic tailgroups are directed upward into the air. However, for this orientation to occur, two conditions must be met: (1) the hydrophobic tail must be long enough to prevent its dissolution in the aqueous media, and (2) the hydrophilicity of the headgroup must be strong enough to prevent the formation of thicker multilayer films at the interface or evaporation of the surfactant molecules. Moving barriers on either side sweep the water surface and force the amphiphiles to pack, forming an ordered, compressed monolayer of the Langmuir film (Figure 5.14).

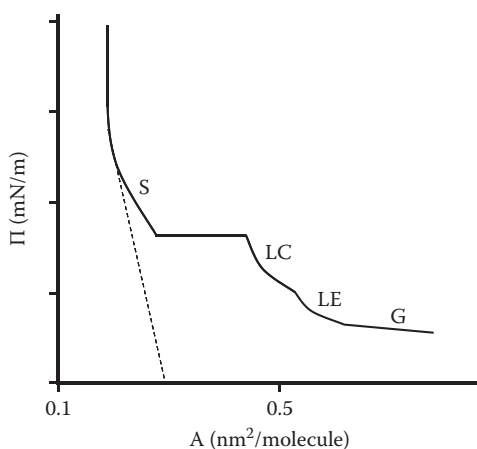


**Figure 5.14** A Langmuir film. Amphiphilic molecules in chloroform solvent are deposited on water surface. The Langmuir film is formed by confining the amphiphiles at the air-water interface with movable barriers.

The packing density of a Langmuir film is crucial in determining the final structure of an LB film. As we have already seen, surface pressure is a measure of the change in the surface tension of a pure liquid due to the presence of a surfactant. Basically, it is defined as the difference between the surface tension of the pure water and the surface tension of the aqueous solution containing amphiphilic solutes. This phenomenon is especially important in Langmuir films. If the number of molecules comprising the Langmuir film is known, then surface pressure as a function of the area occupied by each molecule may be studied. The resulting plot is known as a pressure-area ( $\Pi$ -A) isotherm, and its shape is unique to the molecule used to form the film. Figure 5.15 illustrates a typical isotherm for a long-chain carboxylic acid  $[\text{CH}_3(\text{CH}_2)_n\text{COOH}]$  at the air-water interface. At constant temperature, the  $\Pi$ -A isotherm is given by the changes in surface tension as the monolayer is compressed.

The  $\Pi$ -A isotherm may consist of several regions, although molecules usually exhibit only a few of them. When the molecules are first deposited onto the film, the monolayer behaves most like a 2-D gas (G), shown by Equation 5.14, since no external pressure is applied.

$$\Pi A = k_B T \quad (5.14)$$



**Figure 5.15** A simplified  $\Pi$ -A isotherm for a simple long-chain fatty acid. Extrapolating the slope of the S phase to zero pressure enables one to obtain the molecular area per molecule at zero pressure (see Example 5.2).

In this equation,  $\Pi$  is the surface pressure,  $A$  is the molecular area,  $k_B$  is the Boltzmann constant, and  $T$  is the temperature. In the G phase, all interactions with the water are attractive, so even the hydrophobic tail is in intimate contact with the water. Transition to the liquid-expanded phase (LE) occurs as monolayer compression begins, and the tail starts to lift away from the surface. The transition states form a plateau because the existence of two phases gives the monolayer only one degree of freedom, and constant temperature gives constant pressure,  $\Pi$  (Knobler 1990).

The liquid-condensed phase (LC) forms with further compression. The hydrocarbon chains are highly aligned and a high number of *trans* conformations appear, indicating the achievement of long-range order due to van der Waal's interactions. Finally, application of great external pressure leads to the ordered solid (S) phase, and the area per molecule is comparable to the closely packed chains of the amphiphile's 3-D crystal. Further compression leads to the collapse pressure,  $\Pi_c$ , where the molecules are ejected randomly throughout the film.

Molecular dimensions can be obtained through the  $\Pi$ - $A$  isotherm by extrapolating the slope of the S phase to zero pressure to get the molecular area per molecule at zero pressure,  $A_0$ .  $A_0$  can be compared to theoretical values of the cross-sectional area of hydrocarbon chains, and a close correlation indicates that a closely packed monolayer has been formed with the hydrophobic tails oriented normal to the surface. Furthermore, the  $\Pi$ - $A$  isotherm gives information on the stability of the monolayer at high pressures.

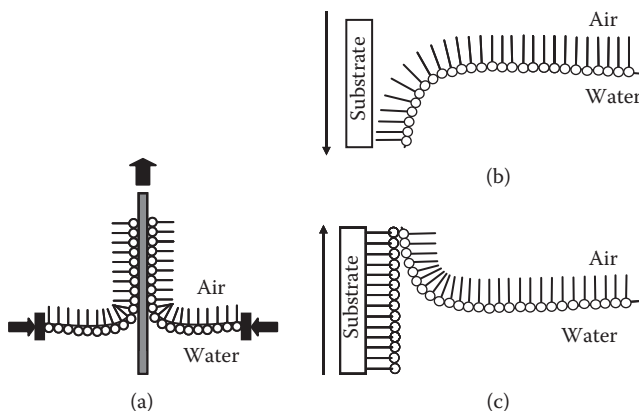
### Example 5.2 Determining the Limiting Surface Area per Molecule

Use Figure 5.15 to determine the area per molecule for the amphiphile described by the isotherm.

*Solution* By extrapolating the slope of the S phase to zero pressure (the dashed line in Figure 5.15), one obtains an area of  $0.2 \text{ nm}^2$  for each molecule.

#### 5.5.2 Langmuir-Blodgett Films

The LB film is subsequently formed by the immersion of a substrate into the water to break the Langmuir film and transfer the monolayer onto the substrate. Additional immersions result in the fabrication of the multilayer film, whose



**Figure 5.16** (a) The LB method is used to transfer a film at the air-water interface to a solid substrate. (b) The downstroke results in the monolayer being transferred to the hydrophobic substrate. (c) Withdrawing the substrate generates a bilayer as a Y-type film.

thickness depends largely on the molecular chain length and number of dippings. The LB deposition technique is illustrated in Figure 5.16.

Deposition of tightly packed molecular monolayers with unilateral molecular directionality normal to the substrate can be induced through annealing. Annealing involves repeated compression and then expansion of the Langmuir film. Studies have found that the calculated area per molecule obtained through the  $\Pi$ -A isotherm actually decreases under such a procedure, leading to a closely packed and well-aligned monolayer.

Film deposition onto the hydrophilic or hydrophobic substrate occurs only when the direction of the motion coincides with the meniscus curvature at the solid-liquid interface. If a hydrophobic substrate is used, the meniscus will curve downward upon immersion into the subphase, resulting in deposition on the downstroke and adhesion of the tails to the hydrophobic surface. If this is the first immersion, subsequent monolayer deposition behavior determines the LB film type. If withdrawal of the substrate results in subsequent layer formation in a head-to-head fashion, then a Y-type LB film is formed. On the other hand, if monolayer transfer occurs only upon insertion of the hydrophobic substrate, then an X-type LB film is formed. A third type of LB film, Z-type, is formed on a hydrophilic substrate where the amphiphilic headgroup adheres to the substrate and transfer occurs only upon withdrawal of the substrate.



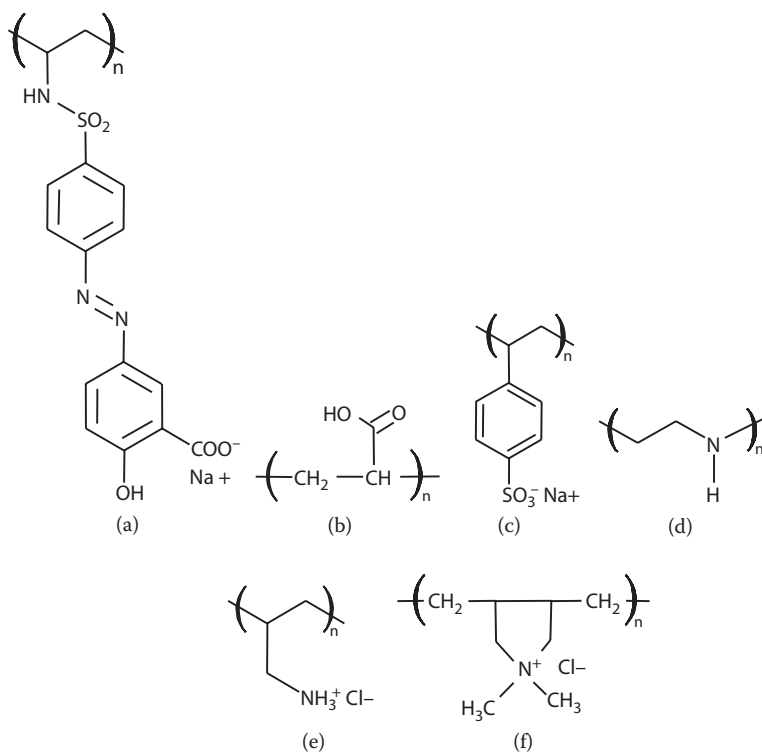
The Y-type film is the most common type of LB film formed. Since monolayer transfer occurs on both the down- and upstrokes, multilayer films may also deposit on hydrophilic substrates with the exception that no monolayer adsorbs in the first insertion. Due to the head-to-head and tail-to-tail conformation, the film formed, if composed of only one kind of amphiphilic molecule, is completely centrosymmetric. Such a film possesses no dipole moment, rendering the assembly useless for NLO applications. However, if fatty acid spacers, such as stearic or arachidic acid, are incorporated between the other amphiphilic species, a net dipole moment may result. This procedure requires two separate LB compartments for each type of molecule. The substrate is inserted through the amphiphile of interest, moved to the other compartment under water, and withdrawn out of the fatty acid spacer Langmuir film. Research has shown a relatively large SHG signal for such a polar-ordered arrangement. Conversely, X- and Z-type films are inherently arranged in an acentric manner with a net dipole moment due to the head-to-tail molecular conformation. In fact, studies by some researchers have found that Z-type LB films generate a strong SHG signal.

In spite of the virtues of the LB technique, it has its shortcomings. One limitation is the unstable nature of the adsorbed films. In practice, although the transfer process occurs very slowly, the layers are not always transferred onto the substrate as would be desired. After deposition, the molecules have been shown to rearrange to more stable conformations, which could lead to loss of order and NLO activity. Another disadvantage is the occurrence of chromophore randomization. It has recently been shown that intermixing between LB layers is present. This produces structural defects and chromophore tilting that becomes amplified when the number of LB layers is increased. Thus, it is not surprising that experimental results show that the SHG signal resulting from the head-to-head, polar-ordered (Y-type) LB films decreased with thickness (Johal et al., 1999). Moreover the fatty acid spacers used for acentricity could lead to even more disorder in the LB film. Stearic acid has been found to form two-dimensional crystalline structures that result in a heterogeneous monolayer filled with defects. Finally, staggered molecular structures may form due to repulsive interaction between strong dipoles. The effect of molecular interactions on the order of the LB film are still sparsely understood, and current research addresses the effects of dipolar interactions on film formation. With the limitations

of the LB technique and other aforementioned methods, the introduction of the electrostatic self-assembly method was readily welcomed.

## 5.6 POLYELECTROLYTES

Polyelectrolytes are polymers, or chains of molecules, that contain free ions that make them electrically conductive. Soluble in water, polyelectrolytes become charged when in solution, and are often countered by a salt ion of opposite charge. Figure 5.17 shows some common polycations and polyanions. The amount of charge on a polyelectrolyte determines whether it is classified as strong or weak. Strong polyelectrolytes are fully soluble, whereas weak polyelectrolytes, with fractional charge, are only partially soluble. Typically,



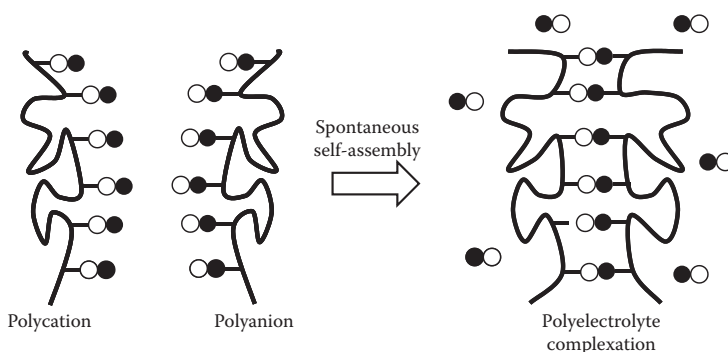
**Figure 5.17** Some common polyelectrolytes. Polyanions are (a) PAZO [poly[1-[4-(3-carboxy-4-hydroxyphenylazo)benzenesulfonamido]-1,2-ethanediyl, sodium salt]], (b) PAA [poly(acrylic acid)], and (c) PSS [poly(styrenesulfonate)]. Polycations are (d) PEI [poly(ethylenimine)], (e) PAH [poly(allylamine hydrochloride)], and (f) PDMA [poly(diallyldimethyl ammonium chloride)].

polyelectrolyte conformation is “flat” because the charges on a linear polymer repel each other due to Coulombic repulsion. With the addition of salt ions, however, this conformation can be altered to a more coiled or collapsed state. This ability to change polyelectrolyte conformation can change their function and is very useful in materials chemistry.

Synthetic organic chemists have been able to control the characteristics of polyelectrolyte assemblies in many different manners: most simply by altering monomeric units, electrolyte group, and polymer length. As a result, polyelectrolytes make a very tunable building block. They have been used in a number of applications ranging from drug delivery agents, to components in conducting films, to facilitating colloidal suspensions, and to creating mimics of biological molecules. Recently, polyelectrolyte microcapsules have found use as microreactors for catalyst activity, precipitation reactions, crystallization reactions, and polymerization reactions.

### 5.6.1 Electrostatic Self-Assembly

One benefit of using polyelectrolytes is that many of them can interact with each other in a process known as electrostatic self-assembly (ESA). As a simple example, Figure 5.18 illustrates the complexation of two oppositely charged polyelectrolytes in the bulk phase. This basic idea can be applied to charged surfaces using polyelectrolytes resulting in the formation of multilayer nanofilms. Because ESA relies on attractive interactions between materials of alternating charge, a number of different kinds of molecules can be used, such as



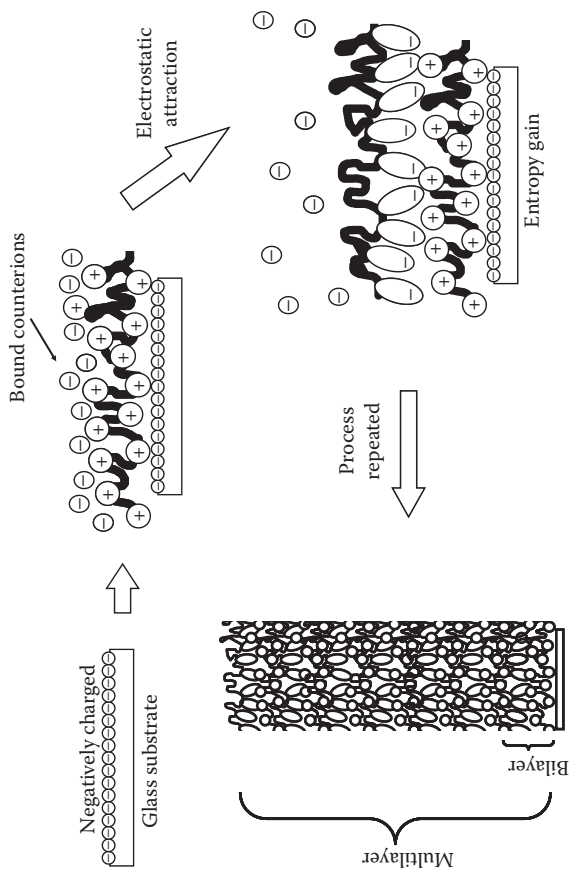
**Figure 5.18** The complexation reaction between oppositely charged polyelectrolytes. The ion-pair formation (or ion exchange reactions) occurs due to a large entropic gain when counterions are expelled into solution. Filled circles represent anionic sites and open circles cationic sites.

small inorganic molecules, proteins, dendrimers, and polyelectrolytes. Our exploration of ESA films will focus on polyelectrolytes. The flexibility of the polyelectrolytes used for ESA, as well as the ease of the self-assembly process, gives it advantages over previous methods of building films. Prior to ESA, the LB technique was the most common for creating films. However, the LB technique requires far more instrumentation, time, and cost to create films that are limited in long-term stability.

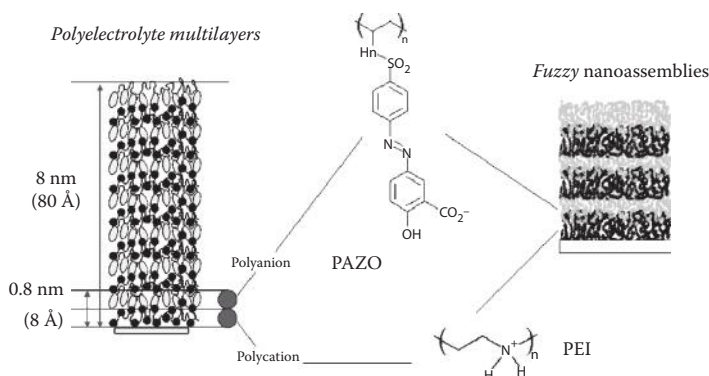
Films formed through ESA techniques can be composed of many adjacent layers and are generally well ordered on the supramolecular scale, depending on the reaction conditions in which they are made. Both the amount of polyelectrolyte as well as the order in which they are deposited can be easily controlled. In the formation of the first layer, a polyion adsorbs to an oppositely charged substrate surface, because of favorable electrostatic attraction. In the case of ESA, polyelectrolytes in aqueous solution adsorb to the substrate surface, since they are removed from solution and irreversibly bound to the substrate by electrostatic interactions. By altering the strength of this attraction, it is possible to alter the amount of polyelectrolyte deposited.

A commonly used system is a positively charged polyelectrolyte on an oxidized silica or silanol surface. The silica substrate surface is negatively charged due to deprotonated silanol groups. When exposed to a cationic solution, the negative charge on the substrate quickly attracts the cations to the surface. ESA methods typically expose a solid substrate surface to the desired adsorbant by “dipping” the substrate into an aqueous solution of the adsorbant for about 10 minutes. In order to build multilayered assemblies, this “dipping” procedure is repeated with alternating cation and anion aqueous solutions, with a rinse step with either water or slightly ionic solution in between that increases the ability of the next layer to be well ordered. The rinse step removes weakly adsorbed molecules that are not a part of the film and primes the surface for the oppositely charged adsorbant. This approach is generally known as the layer-by-layer (LbL) technique.

Repetition of this process can be used to make many layered films. To better understand the overall formation of a multilayer assembly we must break down the mechanisms of ESA further, into adsorption of the polyelectrolyte layer to the substrate film, interpenetration of polyelectrolyte materials within the film, and the complexation of molecules that occurs within the film. Figures 5.19 and 5.20 illustrate how ESA can be used to construct polyelectrolyte multilayer assemblies.



**Figure 5.19** Polyelectrolyte self-assembly. The polycation adsorbs from aqueous solution to the negatively charged substrate. This is followed by exposure of this film to an aqueous solution of the polyanion. Adsorption of the polyanion to the polyanionic film results in a bilayer film. The process can be repeated many times to produce a multilayer nanofilm.



**Figure 5.20** Dimensions of a typical multilayer constructed using the polycation PEI and the polyanion PAZO. The thickness of a bilayer as measured by ellipsometry is around 1 nm. A 10-bilayer film has a thickness between 100 and 150 nm. Interlayer interpenetration is illustrated as the “fuzzy” nanoassembly on the right. The degree of interpenetration is drastically underexaggerated.

### 5.6.2 Charge Reversal and Interpenetration

Within the multilayer film, there are two main regions: the surface, at which charge overcompensation or reversal occurs, and the bulk area, in which interpenetration occurs. The surface of the film is repeatedly exposed to polymer solutions of alternating charge. The bulk is not exposed to the aqueous environment. This greatly impacts the differences in chemistry that occurs in these two main regions.

At the surface, charge reversal, or charge overcompensation, helps drive the formation of layers within the film. Typically, polyelectrolytes are assumed to take a structure with 1:1 charge stoichiometry. However, in order to drive the formation of multilayers, an overall surface charge must be exposed prior to the next deposition step in an oppositely charged polyelectrolyte solution. Charge reversal of the surface enables the electrostatic interactions that drive the self-assembly process.

The charge overcompensated region of the terminal layer is matched by oppositely charged salt counterions in solution. Upon exposure to the next polyionic solution, the counterions are displaced by the polyion. This displacement is driven primarily by entropic gain from displacement of the counterions. Entropic gain combined with electrostatic interactions results in a net negative free energy of adsorption that allows the polyelectrolyte to remain adsorbed to the surface. Further adsorption of the polyelectrolyte will stop when repulsive interactions from material already adsorbed prevail. As a

result, approximately reproducible depositions occur in the formation of each subsequent layer. Once a polyelectrolyte has fully adsorbed to the surface, charge overcompensation occurs again at the terminal layer surface, yielding a reversal in surface charge, compared to the film prior to deposition. As happened for the layer before, the charged surface is balanced by positive counterions from the polyelectrolyte solution. An assembly with one polycationic and one polyanionic layer comprises a single bilayer.

The presence of counterions at the charged surface can be better explained through the use of image forces. Image forces result from the differences in the dielectric constants of water and the film interface. The dielectric constant difference causes a repulsion of water charges from the surface, allowing a polyelectrolyte counterion to bind at the surface instead. These counterions will subsequently be displaced by a polyelectrolyte with the same charge, opposite to the surface—a process driven by gains in entropy because more salt molecules are displaced from the surface than the number of polyelectrolyte molecules that adsorb.

To maintain the overall neutral charge of the film, charge compensation—either intrinsic or extrinsic—must be present. Extrinsic compensation refers to polymer charges that are balanced by extrinsic materials, such as the salt counterions. It has been found that there is actually very little salt concentration observed within the film. Instead, there is a 1:1 stoichiometric ratio of polycation to polyanion charge. As a result, the overall charge balance within the multilayers must be attributed primarily to intrinsic compensation. However, some extrinsic compensation has been found to occur at the substrate surface, where charge overcompensation is neutralized by salt counterions.

Within the bulk of the film, we might expect clearly defined layers of alternating charge. This is not what occurs. Upon adsorption the polyelectrolyte does not simply “lie” on the surface. It has been commonly observed that after adsorption, the outer polyelectrolyte then diffuses slightly into the inner polyelectrolyte layer. This effect results in blurred or “fuzzy” distinctions from layer to layer (Figure 5.20). The effects of interpenetration underscore the lack of organized structure observed within the bulk film.

The degree to which interpenetration occurs, while not completely understood, is linked to ionic concentration and charge density of the polyelectrolyte. In general, ionic concentration is correlated with interfacial overlap. The salt

ions cause conformational changes of the polyelectrolytes by shielding the charges on the polyelectrolytes, enabling a polymer to “extend.” As a result, though it becomes more difficult for the polyelectrolyte to diffuse deeper into the assembly, the presence of salt ions opens up more polymer segments to competitive ion exchange with counterions. Polyelectrolyte charge—whether a polyelectrolyte is strong or weak—also contributes to the extent of overlap observed. With a weaker polyelectrolyte, or a polyelectrolyte with lower charge density, a flat/extended conformation of the polymer cannot be achieved. As a result, the polyelectrolyte cannot penetrate the outer film layer to as great a degree. Overall, ionic concentration plays a significant role (in addition to the charge density of the polyelectrolyte) in contributing to the interpenetration of the polyelectrolyte within the film.

Interpenetration results in the bulk of the film appearing to be more homogeneous, while the surface of the terminal layer has an overall charge. Interfacial overlap also results in a lack of crystalline structuring. This limits the applications of ESA films in producing materials that require ordering for functionality purposes, such as chromophore ordering for nonlinear optical responses. However, the organizational structure of ESA films has been found to be useful for the construction of conductive polymeric materials. The overlap from interpenetration allows an electron to flow easily through the film, which would not be the case if the layers were stratified into organized layers.

### 5.6.3 Multilayer Formation

Charge reversal, interpenetration, and complexation all contribute to the structure and formation of a multilayer. Broader factors to consider for successful multilayer formation are film stability and film thickness. The first, film stability, requires that a completed multilayer assembly be stable in structure when removed from solution, so that the film does not change or degrade following assembly. The second, film thickness, can be tailored based on ESA conditions and number of multilayers.

Film stability requires that the polyelectrolyte assembly be irreversible. As a result, spontaneous desorption of material from the solid substrate film cannot occur in a well-formed assembly. Desorption due to small ion competitors or desorption due to exposure to polyelectrolyte solution all cannot occur in order to maintain the irreversibility of the film. It has been determined that desorption of material occurs at such a low kinetic rate that the ESA film can be considered irreversible.



Control of multilayer thickness is determined by a variety of factors, such as polymer chain length, polymer charge, and strength of ionic solution. The most obvious way to control film thickness is to control the number of multilayers formed on the substrate surface. A greater number of multilayers will result in a thicker overall film, whereas fewer multilayers will decrease the film thickness. In addition to the number of multilayers, layer thickness is highly dependent on ionic concentration, with an almost linear relationship. Nearly linear multilayer growth has been observed, up to 100 bilayers (Reveda and Petkanchin, 1997). Low salt concentrations will allow the polymer to assume a flat, extended conformation. Higher salt concentrations result in greater shielding of the charges on the polyelectrolyte, allowing the charges to be in closer proximity and resulting in a more coiled structure. Low ionic concentration screens charges on the polyelectrolyte and the surface charges, allowing for the flat conformation. The looped and coiled conformation that results from increased ionic concentration creates thicker layers and thus an overall thicker film, though it will not necessarily change the amount of polymer adsorbed to the surface. Charge density on the polymer can also affect film thickness. A lower charge density correlates with thicker layer formation. Similar to the results at high ionic concentration, a polymer with lower charge density will assume a more coiled conformation due to fewer electrostatic interactions between the substrate surface and the weakly charged polyelectrolyte. Perhaps contrary to intuition, it is important to note that polymer molecular weight and polymer branching do not cause significant changes in multilayer thickness.

## **5.7 MODEL PHOSPHOLIPID BILAYER FORMATION AND CHARACTERIZATION**

Biological membranes are complex structures that contain a large amount of proteins and are difficult to study *in vivo*. By using model phospholipid bilayers, we can study how specific membrane components interact by incorporating only these specific components into the membrane (Castellana et al., 2006). Furthermore, phospholipid bilayers closely resemble cell membranes in many ways. They retain two-dimensional fluidity and can be great environments for presenting membrane proteins. Therefore, model membranes have historically been used to study membrane properties *in vitro*. They have also been used for the investigation of biological processes involving

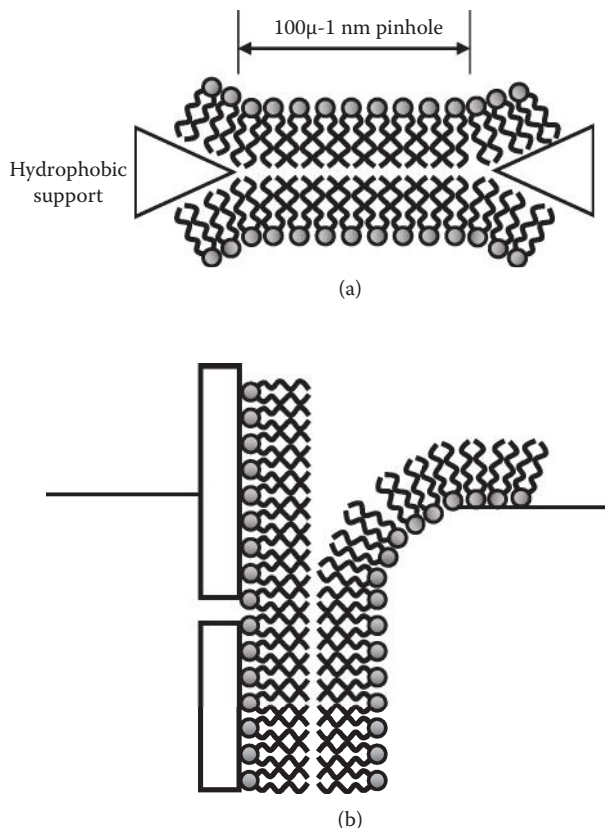
membranes that occur at the cellular level, such as ligand-receptor interactions, viral attack, and cellular signaling events.

A host of techniques are used to study these membrane systems. In this book, discussion will be limited to widely used fluorescence techniques such as fluorescence recovery after photobleaching (FRAP), fluorescence resonance energy transfer (FRET), and fluorescence interference contrast microscopy (FLIC), as well as a few surface-sensitive techniques such as atomic force microscopy (AFM) and quartz crystal microbalance (QCM). There are a number of other important methods, not covered in this section, for studying model bilayers such as electron microscopy, which allows for the direct visualization of membranes *in vivo*, impedance spectroscopy, which allows for electrical measurements of membranes that are important for understanding neurons, as well as neutron and x-ray scattering techniques that are used to probe the structure and periodicity of membranes.

### 5.7.1 Black Lipid Membranes

Several methods for producing black lipid membranes exist. All involve the formation of a membrane over a small aperture usually less than 1 mm in diameter. The hole is formed in a hydrophobic material such as Teflon and is usually part of a wall separating two compartments that can be filled with aqueous solution [see Figure 5.21 (a)]. The two most popular methods for black lipid formation involve the painting of the lipid solution over the aperture or the formation of a folded bilayer. Painting is carried out with a small artist's paintbrush. The formation of folded lipid bilayers requires a container with two compartments separated by a small aperture, and the solution levels in each compartment must be controlled independently [Figure 5.21 (b)]. The desired solution is filled into each compartment and a monolayer of phospholipid material is spread on top of one of the compartments. The solution level containing the monolayer can be raised and lowered over the aperture to deposit the bilayer.

Black lipid membranes have been used to study various biophysical processes. Black lipid membranes are suspended in solution, where there are no unwanted interferences of the membrane with an underlying support. The absence of a support allows for transmembrane proteins to be incorporated within the phospholipid bilayer such that they remain fully mobile and active. The ability to insert single protein pores into black lipid membranes allows for the creation of potential nanodevices. This has been accomplished by Gu et al. (1999) through the use of genetically modified  $\alpha$ -hemolysin.  $\alpha$ -hemolysin



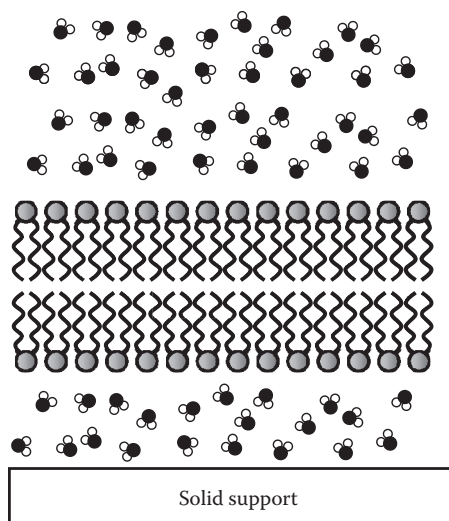
**Figure 5.21** (a) Illustration of a black lipid membrane. (b) The formation of a folded lipid bilayer. (Reprinted from Castellana, E. T. and Cremer, P. S. Solid Supported Lipid Bilayers: From Biophysical Studies to Sensor Design. *Surface Science Reports* 2006, 61.10: 429–444. Copyright 2006, with permission from Elsevier.)

mutants that can non-covalently capture cyclodextrin molecules within their pores were created. A current change is measured due to the restriction of the pore by cyclodextrin. The binding and unbinding of small organic molecules at the single molecule level can be measured using this process.

However, the lack of support limits the lifetime of the bilayer due to poor stability. Furthermore, the methods of detection and characterization of black lipid membranes are also limited. Because there is no support, surface-sensitive techniques cannot be used for characterization.

### 5.7.2 Solid Supported Lipid Bilayers (Richter et al., 2006)

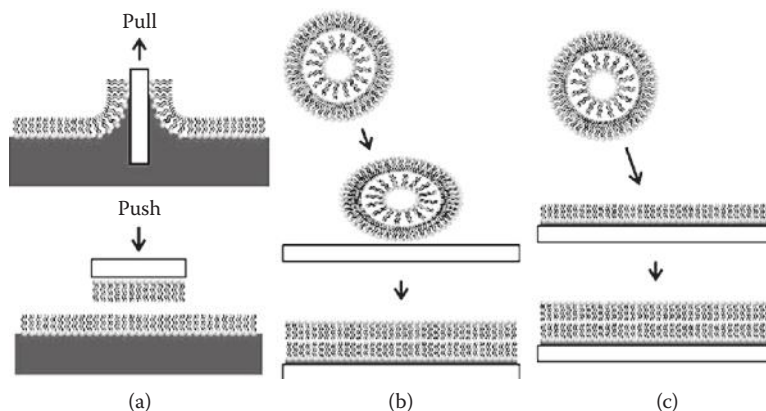
Supported bilayers, as illustrated in Figure 5.22, are more robust and stable than black lipid membranes and can also



**Figure 5.22** Schematic diagram of a supported phospholipid bilayer. (Reprinted from Castellana, E. T. and Cremer, P. S. Solid Supported Lipid Bilayers: From Biophysical Studies to Sensor Design. *Surface Science Reports* 2006, 61.10: 429–444. Copyright 2006, with permission from Elsevier.)

be analyzed by surface-specific analytical techniques such as AFM, QCM, DPI, and SPR. In these systems fluidity is maintained by a 10–20 Å layer of trapped water between the substrate and the bilayer (Groves and Boxer, 2002).

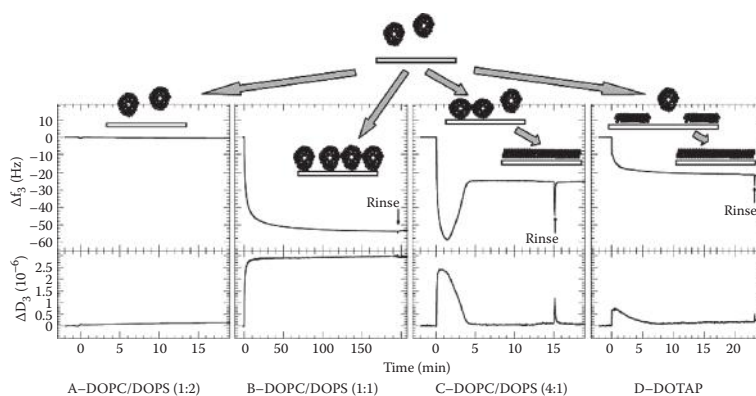
In order to support a high quality membrane the surface should be hydrophilic, smooth, and clean (Tamm and McConnell, 1985). A few commonly used substrates for bilayer formation are silica, mica, gold, and titanium oxide. There are three general methods for the formation of supported phospholipid bilayers on planar supports. The first involves the deposition of a lower leaflet of lipids from the air-water interface by the Langmuir-Blodgett technique [Figure 5.23(a)]. This is followed by the transfer of an upper leaflet by the Langmuir-Schaefer procedure, which involves horizontally dipping the substrate to create the second layer. A second method of supported bilayer formation is the adsorption and fusion of vesicles from an aqueous suspension to the substrate surface [Figure 5.23(b)]. The third method, a combination of the first two methods, can be employed by first depositing a monolayer via the Langmuir-Blodgett technique followed by vesicle fusion to form the upper layer [Figure 5.23(c)]. Each of the three methods contains particular advantages and disadvantages. The first and third methods are useful for the formation of asymmetric bilayers.



**Figure 5.23** Common methods for forming supported bilayers. (a) Langmuir-Blodgett technique to deposit monolayer followed by pushing the substrate horizontally through another lipid monolayer. (b) Vesicles in solution adsorb and fuse to the surface to form a bilayer. (c) A combination of (a) and (b). (Reprinted from Castellana, E. T. and Cremer, P. S. Solid Supported Lipid Bilayers: From Biophysical Studies to Sensor Design. *Surface Science Reports* 2006, 61.10: 429–444. Copyright 2006, with permission from Elsevier.)

The adsorption and fusion of small unilamellar vesicles (SUVs) is one of the easiest and most versatile means for forming supported bilayers. Small unilamellar vesicles can be prepared many different ways. Lipids are often stored in chloroform to prevent degradation. After a known quantity of lipids is withdrawn from the chloroform suspension, they should be placed into a small glass container with a large surface area. The chloroform can then be evaporated with gaseous nitrogen. This makes a thin coating of lipids in the container, which should be dried in a vacuum manifold to ensure that there is no residual chloroform. The dried lipids can be rehydrated to multilamellar vesicles by exposing the film to whatever solvent is desired to be trapped within the vesicle, then sonicating them. Extrusion of these multilamellar vesicles through porous polycarbonate membranes at high pressure will result in SUVs with a size distribution dependent on the size of the polycarbonate membrane's pores.

Parts of the mechanism of vesicle adsorption and fusion to form a bilayer have been elucidated by QCM studies (Figure 5.24). The process begins with the adsorption of SUVs from the bulk solution onto the substrate. In the early stages after adsorption, SUVs may fuse with one another to form larger unilamellar vesicles. The vesicles then rupture, forming a supported bilayer in a process that depends upon the destabilizing interactions affecting the vesicle, such as the



**Figure 5.24** Lipid deposition pathways measured by QCM-D on silica. (a) Vesicles do not adsorb. (b) Vesicles adsorb and remain intact, forming a supported vesicular layer (SVL). (c) Vesicles adsorb and remain initially intact. At high vesicular coverage an SLB is formed. (d) Vesicles adsorb and rupture instantaneously to form an SLB. (Reprinted with permission from Richter, R. P., Berat, R., and Brisson, A. R. Formation of Solid-Supported Lipid Bilayers: An Integrated View. *Langmuir* 2006, 22.8 : 3497–3505. © American Chemical Society.)

osmotic pressure due to a salt concentration within the vesicle that differs from the salt concentration outside the vesicle, or the attraction between the individual lipids and substrate that causes vesicles to deform from their more stable spherical shape into a more oblong shape. Figure 5.24 shows four ways that lipid vesicles interact with the solid support as followed by quartz crystal microbalance with dissipation monitoring (QCM-D). The QCM-D technique is a valuable tool for screening the overall properties of the deposited lipids. The dissipation parameter allows for distinguishing between intact, adsorbed vesicles (high dissipation) and bilayer patches (low dissipation). As shown in Figure 5.24, vesicles either do not adsorb [Figure 5.24(a)]; adsorb and remain intact, giving rise to a supported vesicular layer [Figure 5.24(b)]; or form a supported bilayer [Figure 5.24(c) and (d)]. As observed by QCM, supported bilayer formation can occur via two scenarios with distinct kinetics. In one case the vesicles rupture quickly upon interaction with the solid support [Figure 5.24(d)]. In another case, a large amount of intact vesicles are adsorbed at an intermediate stage of the process [Figure 5.24(c)].

The main disadvantage of supported bilayers is that the supported bilayer is not truly decoupled from the underlying substrate. The layer of hydration at the bilayer-support interface allows for membrane fluidity, but this layer is also too thin to prevent transmembrane proteins from interacting

unfavorably with the underlying substrate. Such interactions can denature the protein or cause the membrane to lose its fluidity.

### 5.7.3 Polymer Cushioned Phospholipid Bilayers

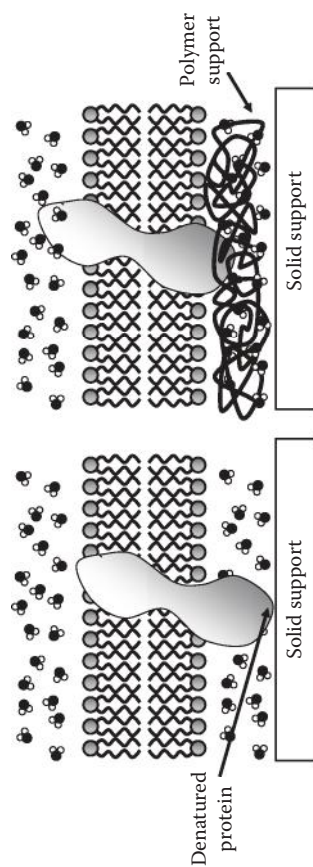
Figure 5.25 also shows the same system in the presence of a lipopolymer support that prevents denaturation. The addition of a polymer layer effectively decouples the membrane from the surface and still allows for investigation by an array of surface science techniques. A well-designed polymer cushion should behave much like a cytoskeleton, present in eukaryotic cells. In physisorbed systems, weak interactions between the phospholipid bilayer and the polymer support can result in an unstable system. This can be overcome by covalently attaching the polymer to the substrate. Next, anchor lipids or alkyl side chains capable of inserting into the phospholipid bilayer are employed in order to further increase stability. These effectively tether the membrane to the underlying support and allow for the incorporation of membrane proteins without denaturing them. Unfortunately, this ability to study membrane proteins comes at the cost of membrane fluidity. In general, it is desirable for the polymer cushion to be soft, hydrophilic, and not too strongly charged.

Two classes of polymers, polyelectrolytes and lipopolymers, are emerging as popular choices for cushion material. Polyelectrolyte cushions can be directly adsorbed from solution to a variety of substrates by means of layer-by-layer deposition. Lipopolymers consist of a soft hydrophilic polymer layer presenting lipid-like molecules at their surface that can insert into a phospholipid membrane and tether it to the polymer spacing.

### 5.7.4 Fluorescence Recovery after Photobleaching

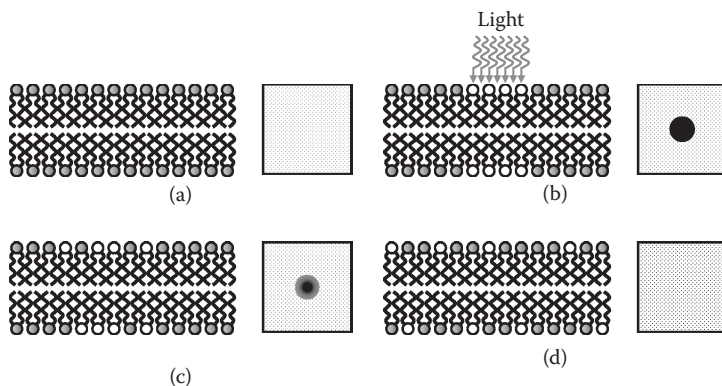
FRAP is an optical technique capable of quantifying the two-dimensional lateral diffusion of fluorescently labeled lipids in a bilayer. This technique is one of the most common techniques performed in order to characterize lipid bilayers. The technique begins by saving a background image of the fluorescing bilayer. Next, a light source is focused onto a small patch of the viewable area either by switching to a higher magnification objective or shrinking the field of view via a pinhole. The fluorophores in this region are quickly photobleached, which is the photochemical destruction of a fluorophore, and an image reveals a noticeable dark spot surrounded by undamaged fluorophores in the bilayer [Figure 5.26(b)]. As





**Figure 5.25** Peripheral domains of transmembrane proteins can become immobilized and denatured on a solid support. A polymer cushion can help shield the protein from the substrate. (Reprinted from Castellana, E. T. and Cremer, P. S. Solid Supported Lipid Bilayers: From Biophysical Studies to Sensor Design. *Surface Science Reports* 2006, 61.10: 429–444. Copyright 2006, with permission from Elsevier.)





**Figure 5.26** Principle of FRAP. (a) The bilayer is uniformly labeled with a fluorescent tag. (b) The label is selectively photobleached. (c) The intensity within the photobleached region is monitored as a function of time. (d) Eventually uniform intensity is restored.

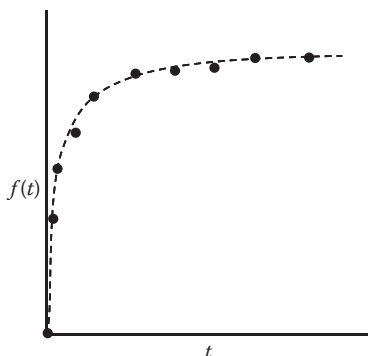
Brownian motion proceeds, the still-fluorescent lipids will diffuse throughout the sample [Figure 5.26(c)] and replace the destroyed lipids in the bleached region. This diffusion proceeds in an ordered fashion, which can be modeled by the diffusion equation (Smith et al., 2008), which shows how the fluorescence intensity  $f(t)$  recovers (increases) over time,  $t$ :

$$f(t) = e^{-2\tau_D/t} \left[ I_0 \left( \frac{2\tau_D}{t} \right) + I_1 \left( \frac{2\tau_D}{t} \right) \right] \quad (5.15)$$

where  $I_0$  and  $I_1$  are known as modified Bessel functions, and  $\tau_D = r^2/4D$  where  $r$  is the radius of the bleached area at  $t = 0$  and  $D$  is the diffusion coefficient. The diffusion constant  $D$  can be simply calculated from a fit of this equation. Diffusion constant values for lipids tend to be in the range of 0.5–5  $\mu\text{m}^2/\text{s}$ . Figure 5.27 shows what a typical FRAP profile looks like (a plot of fluorescence intensity versus time). Fitting this curve to Equation 5.15 allows the determination of  $D$ .

### 5.7.5 Fluorescence Resonant Energy Transfer

FRET is a useful tool for quantifying molecular dynamics in nanoassemblies of biophysical and biochemical importance. For monitoring the locations of two molecules relative to one another, one of them is labeled with a donor and the other with an acceptor fluorophore. When they are not near each other, only the donor emission is detected upon donor excitation. However, when the donor and acceptor are in proximity (1–10 nm) due to the interaction of the two molecules, the acceptor



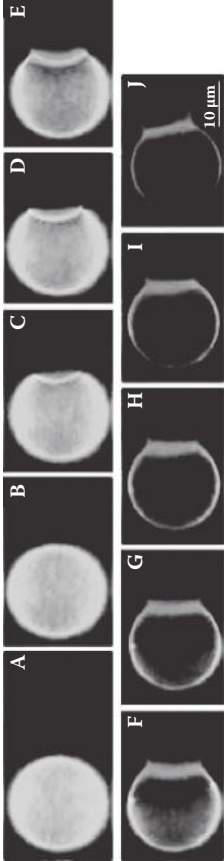
**Figure 5.27** Illustration of a typical FRAP recover curve with corresponding images. The solid line represents a fit to the data using a 2D diffusion model equation.

emission is predominantly observed because of the intermolecular transfer of energy from the donor to the acceptor.

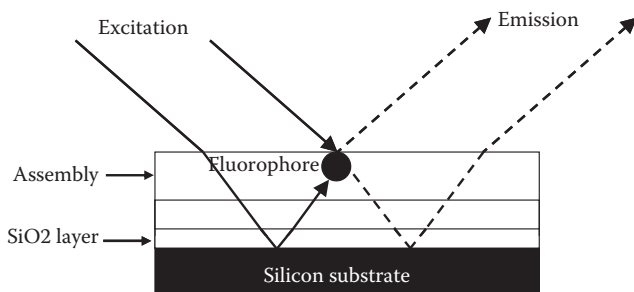
FRET has been used by Lei and Macdonald (2003) in order to monitor the fusion of two lipid vesicles. Video images of two vesicles not fusing and two vesicles fusing are shown in Figure 5.28. As the vesicles come into contact, the contact area becomes a different color (depending on the fluorophores used—red and green in this case), corresponding to simultaneous emission of both red and green from the same area (shown in greyscale in the figure). Thereafter there are dramatic color and intensity changes [Figure 5.28(j)], which are interpreted as hemifusion; the color of the contact area changes, the contact area becomes slightly fatter, and then red color can be seen beginning to diffuse from the contact area over the surface of the positive vesicle, reducing the intensity of the green color in the process. This donor-acceptor relationship was examined in more detail in Section 5.1.

### 5.7.6 Fluorescence Interference Contrast Microscopy

In FLIC, the proximity of fluorescent probes to a reflective plane leads to the modulation of the fluorescence intensity, analysis of which can provide nanometer-scale topographic information. FLIC can have resolution of a few to hundreds of nanometers. The principle underlying FLIC is that in the presence of a reflective surface, light can travel to a fluorophore along a direct path or a reflected path; the difference in optical path length between these two routes leads to interference (Figure 5.29). The resulting intensity is a function of this path-length difference, which is a function of the height of the fluorophore above the reflective plane. Interference occurs independently



**Figure 5.28** FRAP images showing the fusion of two vesicles. The red and green color is shown in grayscale. (Reprinted from Lei, G. and MacDonald, R. *Biophysical Journal*, 2003, 85, 1585–1599, Biophysical Society, Elsevier. Copyright 2003, with permission from Elsevier. See Lei and MacDonald for the color image.)



**Figure 5.29** Basic principles of a FLIC experiment. The resulting intensity is a function of the path length difference between the direct and reflected light paths, which is a function of the height of the fluorophore above the reflective plane.

for the excitation and the emission light, each of which must be considered separately. The typical setup makes use of a silicon wafer as the reflector. A transparent oxide ( $\text{SiO}_2$ ) layer acts as a spacer. The system of interest (i.e., a lipid bilayer with fluorescent probes) is placed on top of the oxide layer.

FLIC can also be used to look at the topography of a secondary membrane. A secondary lipid bilayer can be assembled atop a solid supported bilayer via the rupture of giant lipid vesicles with diameters that approach tens of microns. The secondary membrane is separated from the primary membrane by a confined layer of water and is therefore free to exhibit nanometer-scale height fluctuations.

## 5.8 SELF-ASSEMBLED MONOLAYERS

A self-assembled monolayer is formed when a particular group on a molecule has a strong affinity for a specific surface. Thiols and silanes are examples of such molecules, and these are discussed in the next two sections. Spontaneous chemisorptions of these adsorbate molecules may occur from either the vapor or liquid phase and often leads to the formation of an organized monolayer. The adsorbate, being a self-assembling molecular building block, often has a relatively simple structure with the reacting group attached to an alkyl chain and terminated by some other inert group. The chemisorption step is usually rapid, lowering the surface energy of the substrate. This is followed by slow two-dimensional organization of the alkyl chains. This organization step is dominated by inter-chain hydrophobic interactions.

Surface organization itself typically occurs in stages, where each stage represents a distinct phase. First, a low-density

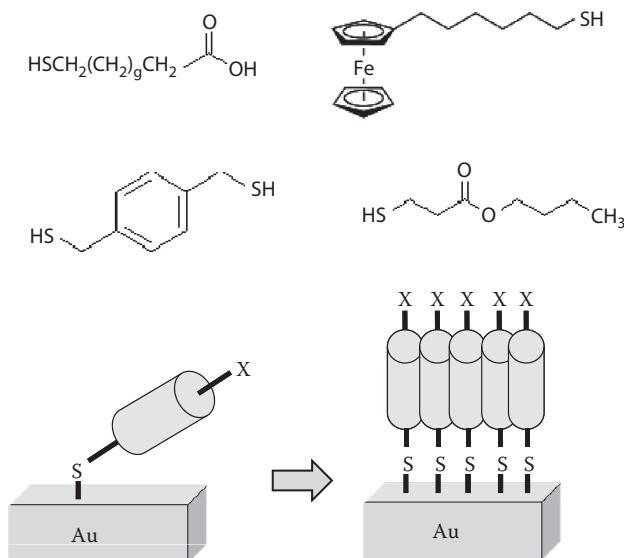
phase is formed comprised of adsorbate molecules randomly dispersed on the surface of the substrate. This is followed by a transition to a second intermediate phase with all of these randomly dispersed molecules lying flat on the substrate. The final transition to a high-density phase involves the molecules orienting themselves normal to the substrate surface. Exactly how these phase transitions proceed depends on temperature. If lateral interactions are ignored, chemisorption follows the Langmuir adsorption isotherm. The kinetics of adsorption is approximately described by Equation 5.16:

$$d\theta/dt = k(1 - \theta) \quad (5.16)$$

where  $\theta$  is proportional to the amount of area occupied and  $k$  is the rate constant.

### 5.8.1 Thiols on Gold

Thiols are compounds having the structure RSH (where  $R \neq H$ ). These molecules are also referred to as mercaptans. Some commercially available thiols are shown in Figure 5.30. The sulfur atom of the thiol is extremely reactive to gold and other noble metals resulting in the formation of a strong metal-sulfur covalent bond having a bond strength on the order of  $100 \text{ kJ mol}^{-1}$ .



**Figure 5.30** Upper: Some examples of thiols used to functionalize a gold surface. Lower: Thiol bound to a gold substrate via a strong Au-S bond. The grey tube is typically a hydrocarbon chain of specified length.  $X$  represents some surface functional group.

The resulting monolayer on the metal surface is thermally stable and is resistant to various solvents and electrolytes.

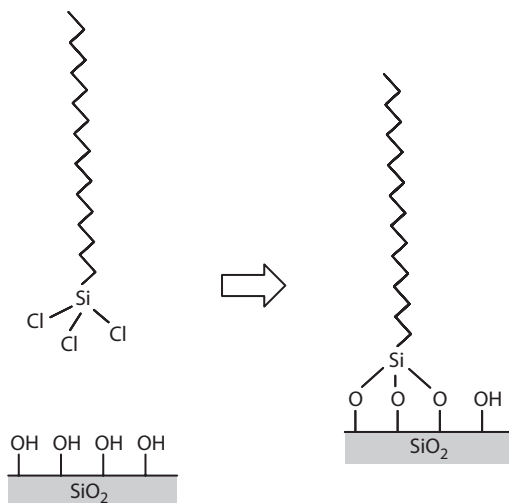
Gold is usually the preferred metal for thiol deposition, mainly because it is inert, biocompatible, and can withstand harsh chemical cleaning treatments. Gold is also easy to pattern using lithographic methods. Low molecular weight thiols can be deposited on gold surfaces by vapor deposition. Most other thiols can be deposited using a simpler method. First the thiols are dissolved in an organic solvent such as chloroform or ethanol. The gold is then immersed in this dilute solution (~1 mM) for 12 to 72 hours at room temperature. The metal is then removed, washed with appropriate solvents, and dried in a stream of nitrogen. It should be stressed that the gold surface needs to be free of impurities before being immersed in the thiol solution. The surface is usually cleaned in appropriate solvents and then exposed to intense UV light in the presence of oxygen for about 30 minutes. The production of oxygen atoms and ozone during the UV exposure causes organic impurities on the gold surface to be oxidized.

The final structure of SAMs constructed using thiols depends on the curvature of the substrate. Planar substrates have no curvature. However, SAMs on nanoparticles such as colloidal particles and nanocrystals tend to stabilize the reactive surface of the particle and provide specific organic functional groups at the particle-solvent interface. This surface functionalization is particularly useful for applications, such as immunoassays, that are dependent on the chemical composition of the surface.

### 5.8.2 Silanes on Glass

Silanes are generally used to construct SAMs on nonmetallic oxide surfaces, such as those found on silicon and glass surfaces ( $\text{SiO}_2$ ). The most commonly used silane to produce a SAM is octadecyltrichlorosilane (also known as OTS). The structure of OTS is shown in Figure 5.31. It is an organometallic compound in which a long octadecyl chain is connected to the reactive trichlorosilane group ( $\text{R-SiCl}_3$ ). OTS reacts violently with water and is sensitive to air. Like thiols, the substrates (glass) are immersed in a dilute solution of OTS, where the solvent is organic.

OTS reacts with the Si-OH groups found on the surface of clean glass or silicon. The reaction is illustrated in Figure 5.31. In order for the reaction to proceed efficiently, the glass substrates must be cleaned thoroughly. This is usually achieved by first immersing the glass slides into a hot mixture of



**Figure 5.31** Octadecyltrichlorosilane (or OTS) chemisorbed on a clean glass substrate resulting in a close-packed SAM.

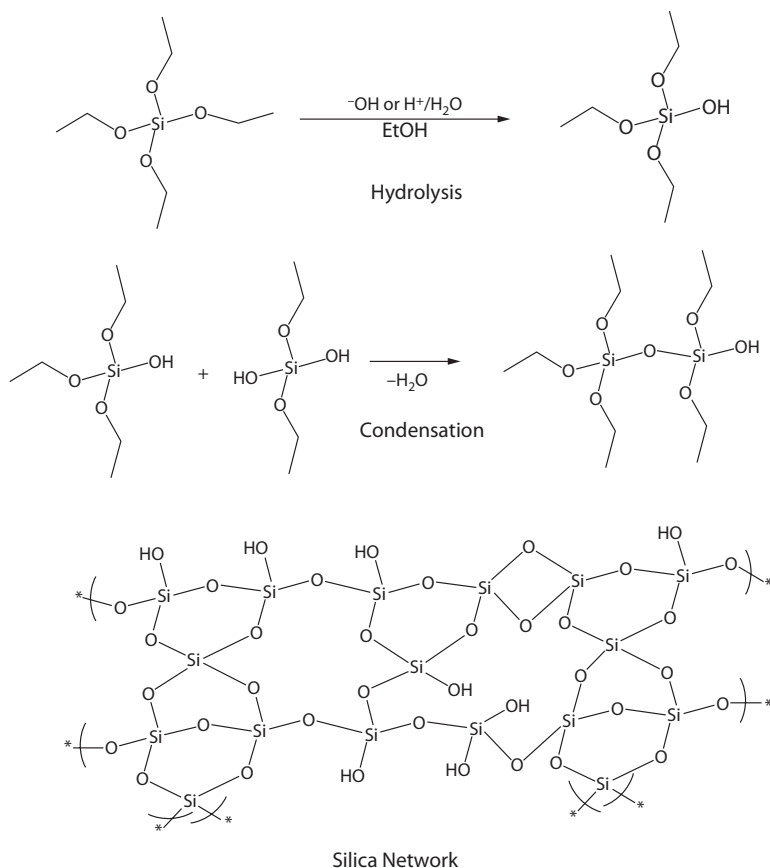
concentrated sulfuric acid and 30% hydrogen peroxide for about 30 minutes. This mixture, known as piranha etch solution, is a powerful and extremely hazardous oxidizing agent that rapidly removes organic impurities from the glass surface. Silicon is usually cleaned by exposing the surface to high intensity UV light for about 30 minutes. As mentioned previously, the production of oxygen atoms and ozone causes organic impurities on the silicon to be oxidized.

OTS and other silanes such as dodecyltrichlorosilane are used in the semiconductor industry to form nanofilm SAMs on silicon dioxide substrates. More specifically, these molecules act as thin insulating gates in metal-insulator semiconductors. OTS is also used in conjunction with conducting polymers in organic-substrate LCD displays.

An interesting extension of the silane chemistry discussed above is the sol-gel processing method used to prepare highly porous silica networks. As an example, let's consider ethoxysilane (Figure 5.32), which can easily be hydrolyzed to the  $(\text{R-O})_3\text{Si-OH}$  form. Since silanes are highly reactive to OH groups,  $(\text{R-O})_3\text{Si-OH}$  molecules can self-react in a condensation process leading to the silica network containing nanopores.

## 5.9 PATTERNING

As the field of nanotechnology expands, the demand for surface manufacturing techniques that are cheaper, more flexible,



**Figure 5.32** The sol-gel process. Silanes are highly reactive to OH groups and  $(R-O)_3Si-OH$  molecules, and can self-react in a condensation process leading to the silica network containing nanopores.

and allow greater feature resolution will continue to increase. The word “nanolithography” can refer to a wide variety of nanoscale surface manufacturing techniques for creating patterned surfaces, such as those found in semiconductor circuits. There are many different ways that such patterns can be fabricated. Here we present a survey of these methods, their advantages, their disadvantages, and possible applications.

### 5.9.1 Optical Lithography

One important method is optical lithography (or photolithography), which is a method capable of producing patterns less than 100 nm using very short wavelength ( $\sim 190$  nm) lasers. The wavelength of light determines the resolution of the



nanopattern. For example, x-ray lithography can reach a resolution of ~15 nm by using light of wavelength ~1 nm.

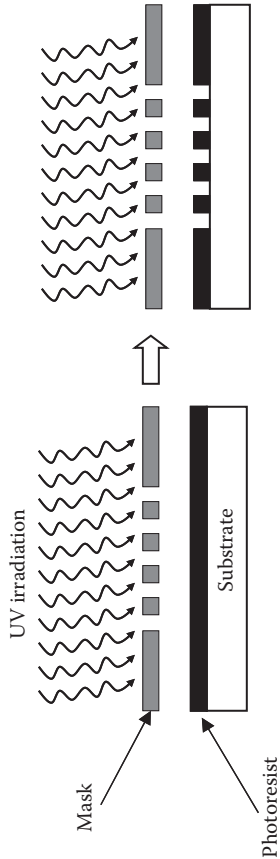
Contact lithography is a common form of photolithography. A photomask is placed over the region to be patterned. The region is typically a flat substrate coated with an imaging photoresist layer, as shown in Figure 5.33. A photoresist is essentially a light-sensitive material that will typically degrade or “burn away” after being exposed to light. The general concept consists of aiming electromagnetic radiation at a surface through openings in a mask template that protects the surface so that only areas not covered by the mask are exposed. Then a chemical that functions as a developing solution is used to remove either the exposed or unexposed areas, leaving either a “negative” or “positive” imprint of the mask. In this way photolithography selectively removes parts of a thin film or the bulk of a substrate. The procedure is usually followed by a chemical treatment to either remove the degraded material or to engrave the exposure pattern into the material beneath the photoresist. The technique excels at creating regular, patterned surfaces, and the process produces surfaces fairly quickly, but the fabrication of the mask can be difficult.

As mentioned above, the resolution of the pattern is determined by the wavelength of light. Usually deep UV light is used to create the pattern down to about 50 nm. Equation 5.17 shows how the minimum feature size ( $\alpha$ ) is related to the wavelength ( $\lambda$ ) of light used.

$$\alpha = k \frac{\lambda}{\phi} \quad (5.17)$$

The constant  $k$  in Equation 5.17 is usually around 0.4, and  $\phi$  is the numerical aperture of the lens seen from the substrate. This equation tells us that the resolution can be improved by using a smaller wavelength and a larger numerical aperture. Maximum feature resolution,  $d$ , as a function of wavelength is  $\lambda/2$ . The lower limit for  $d$  is around 100 nm. However, resolutions under 100 nm have been achieved by using the principles of interference combined with photolithographic methods.

Various types of radiation can be used depending on one's purpose. Electron beams in particular can create surface structures with good resolution, down to 116 nm in height or depth and 105 nm in width. Additionally, electron beams have been used to create these structures on a silicon surface without intermediate developing stages; the beams themselves etch the surface. Electron-beam lithography suffers the



**Figure 5.33**    Patterning a substrate using UV light. The light passes through a photomask and degrades exposed regions of the photoresist.

drawback of not being very up-scalable, and it is also time and money intensive.

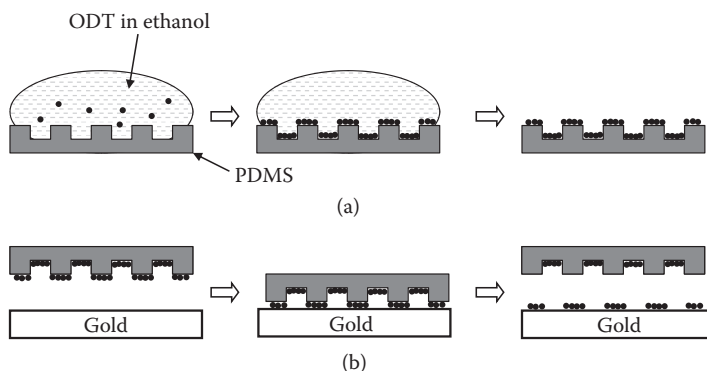
### 5.9.2 Soft Lithography

Another common patterning technique is soft lithography. This method refers to a family of lithographic techniques for patterning using molds and elastomeric stamps. The stamps are usually pieces of polydimethylsiloxane (PDMS) that have been patterned usually against a master to form a relief pattern.

As an example, let's see how a pattern of octadecane thiol (ODT) molecules can be transferred to a gold substrate (Figure 5.34). First the stamp is placed in a solution containing ODT. ODT molecules adsorb onto the PDMS surface forming the "ink" component of the stamp. The solvent is removed and the PDMS stamp is put in contact with a gold substrate, where the ODT molecules spontaneously chemisorb onto the gold surface. Thus the pattern from the stamp is transferred for the gold via the ODT ink. Further details on the chemisorption of thiols onto gold can be found earlier in this chapter.

### 5.9.3 Nanosphere Lithography

Nanosphere or colloidal lithography uses a packed array of nanospheres as the mask for the surface to be patterned. Chemicals can be deposited in the exposed areas between the spheres, or the surface in these exposed areas can be etched by using ion beams. The pattern produced can be manipulated in several ways. For example, chemical deposition can be done



**Figure 5.34** (a) A solution of ODT is placed over the PDMS stamp. The solvent is removed and ODT molecules assemble on the stamp. (b) The PDMS stamp with the ODT "ink" is placed on a gold substrate. When the stamp is removed, the ODT molecules in contact with the gold surface chemisorb to the substrate, transferring the pattern from the stamp to the gold via the ODT ink.

at different angles, or heating the packed array of nanospheres at different temperatures can cause it to deform in desired ways. It is also possible to create multiple layers of sphere arrays, where the upper layers can act as a mask for the lower layers, leaving a complex three-dimensional structure. It is worth noting that the resolution is dependent on the size of the nanospheres used in the mask. By using sufficiently small spheres, one can achieve resolutions below 100 nm.

This technique has the benefit of being simple and inexpensive compared to other nanolithographic techniques. Nanosphere lithography can also produce structures that are difficult or impossible to create with other nanolithographic techniques. Unfortunately, it is difficult to produce a flawless surface with the method; errors occur as the result of difficult-to-avoid disorder in the nanosphere mask, which may lead to small differences between masks that are formed in similar conditions. Furthermore, the space between features and their size cannot be changed independently, as both are directly determined by the size of the nanospheres used. The number of structures that can be created by this method is also limited, because structures created by nanosphere lithography must be based on the unoccupied interstices of a packed sphere array. For example, it is hard to imagine a field of square nanosphere pillars being produced by a mask of patterned spheres.

#### **5.9.4 Patterning Using AFM**

In addition to being used as an imaging tool, the atomic force microscope (AFM) can be used to manipulate surfaces on a very small scale. In standard use, the AFM images a surface in three dimensions by measuring the deflection of a small, very sharp tip as it moves along a surface. For patterning, however, the tip can be used to interact with a surface and alter it in order to produce very small structures.

Some nanolithographic techniques that employ AFM may be termed “destructive”—they etch or otherwise damage a surface to create the intended surface structures. The tip can be used to break apart surfaces on a very small scale, or to move very small objects around on a surface. The AFM tip can also be modified to irreversibly alter the surface in various ways. For example, a catalyst can be added to the tip to initiate reactions. Another common method for surface alteration is the initiation of very localized redox reactions by running a current through a conductive AFM tip. UV light can be used to control this localized redox process because it affects electron-hole pairs in a way that changes their resistance to current

flow. This means that a low-voltage current running through the tip will be able to oxidize a surface only when UV light is shined on the surface at the same time. This method can produce smooth surface features with a height of about 40 nm and a lateral length of about 1  $\mu\text{m}$ . Surface-destructive AFM is thus fairly flexible, but its use can be limited by expense and low production speed.

Other nanolithographic techniques that use AFM may be termed “constructive” because they build up the intended structure on the surface. As with destructive AFM nanolithography, there are several different possible implementations. A common constructive method is the use of AFM as a “dip pen.” The AFM tip is covered in some molecule to be used as “ink,” which is then deposited on the surface at the meniscus between the tip and the surface. A resolution below 50 nm is currently possible with dip-pen methods.

Metals can be transferred to the tip by applying pressure to the tip on a metal surface, heating it, and allowing it to cool. The substance can be transferred from the tip to a surface by reheating the tip as it is dragged over the surface. Using this technique with indium has been successful in the creation of conductive nanoscale wires, with widths as low as 50 nm.

There are many different molecules that can be used as ink for this method, giving a great deal of flexibility in its uses for surface patterning. For example, a surface can be patterned with chemicals that will protect it against a later etching treatment, allowing for an interesting combination of constructive and destructive lithographic techniques to make more complex patterns. Another possible use of this method is the patterned deposition of particles that can act as “seeds” for new pattern creation, either by catalyzing reactions or by binding to other molecules with interesting shapes that form the basis for a new pattern.

An advantage of dip-pen nanolithography is that the technique can also be scaled up by using an array of several AFM tips instead of just one, allowing something analogous to desktop printer functionality for nanofabrication. Using an array of tips would allow for the creation of patterns at greater speeds and over larger areas. This has shown great potential for the use of AFM in the creation of nanoscale electronic components such as transistors.

Dip-pen techniques also have drawbacks, the most unfortunate of which is that the created patterns are not perfectly reproducible. A number of different factors affect molecule deposition: temperature, humidity, the type of molecule used

as ink, the characteristics of the AFM tip, the write speed, and the pressure used can all affect the pattern ultimately produced on the surface. For example, altering the humidity can change the size and structure of the meniscus where ink is deposited, resulting in a larger or smaller area of the surface onto which the same amount of ink should have been deposited.

“Fountain-pen” methods, in which the tip is placed within a nanoscale tube that delivers a constant flow of “ink” molecules to the tip, are a related surface fabrication method with similar advantages and drawbacks. The resolution here is worse than dip-pen methods, as the fountain-pen methods can produce only patterns with feature widths above 200 nm.

AFM tip hammering nanolithography is yet another variation on AFM nanolithography. In AFM tip hammering nanolithography, the AFM tip is used to imprint the surface. One paper by Wang et al. (2009) gives an example of the technique being used on a polystyrene-*block*-poly(ethylene/butylenes)-*block*-polystyrene (SEBS) copolymer. The AFM tip deforms the polystyrene sphere component of the copolymer, resulting in an imprinted area. Embossing as well as imprinting is possible; one simply imprints the area around the area to be embossed to get a raised structure.

This technique has several advantages. For example, the surface imprint is reversible. Depending on the copolymer used as the surface, heating it at high temperatures that allow for reorganization can erase the imprint. This does not mean that the imprint is unstable at all temperatures. At room temperature, below the threshold energy required for reorganization of many molecules, the patterns are fairly stable. For the SEBS copolymer, it was found that the imprint is erased in about 5 minutes at 70°C and that around 50% of the original surface contrast remains after 70 days at 25°C. More stable copolymers would require a higher temperature to reverse the patterning and clear the imprint, but they would likely also be more stable for longer at room temperature. Also, the AFM tip hammering nanolithography has a very good feature resolution and is currently able to produce patterns with features as small as 13 nm for imprinting and 18 nm for embossing. The combination of feature resolution, surface stability, and reversibility makes the method attractive for high-density data storage. The method also has drawbacks, the primary one being surface specificity. The surface chosen must be imprintable with a reasonable level of force, but also stable enough that the imprinting will survive a desirable amount of time.

### 5.9.5 Summary

Nanolithographic methods can be used individually or combined with other fabrication methods to amplify the range of possible surface structures and their uses. For example, photolithography can be used to create a pattern on a surface on which nanowires can be grown. The result is multiple levels of roughness, a surface property with practical applications in superhydrophobicity. Nanolithography is thus a quickly developing field encompassing many techniques for the fabrication of very useful nanoscale patterned surfaces.

## 5.10 DNA AND LIPID MICROARRAYS

Over the past two decades, advances made in our understanding of genetics have led to the creation of DNA microarrays. DNA microarrays are surfaces that have been coated with specific oligonucleotide sequences and have been particularly useful in gene expression studies. DNA microarrays vary in the number of oligonucleotides as is determined by the nature of the study; diagnostic DNA microarrays generally use tens of oligonucleotides, whereas those for research and screening can have hundreds of thousands of oligonucleotides on a single microarray. This section will address the basic protocol of a DNA microarray experiment, the various methods of array fabrication, optimization of an array, and their applications. The student is encouraged to review basic biochemistry of DNA before reading this section.

### 5.10.1 Using a DNA Microarray

Each DNA microarray is unique to a certain purpose, such as diagnosing cardiovascular disease or monitoring the entire genome of a strain of *Escherichia coli*; as such, a microarray must be fabricated or purchased for each experiment. Depending on the size of the microarray and the availability of equipment, it is often more feasible to purchase prefabricated microarrays. Fabrication techniques will be described in more detail later in this section.

Each spot on a DNA microarray contains thousands of oligonucleotide molecules called probes that complement their corresponding mRNA. A sample is prepared for DNA microarray analysis by extracting the cells' total RNA. The RNA can be purified by treating it with poly(A) polymerase followed by oligo(dT) chromatography. The purified RNA is then labeled for detection; in the past,  $^{33}\text{P}$  has been used, although radioactive labeling has mostly been replaced by

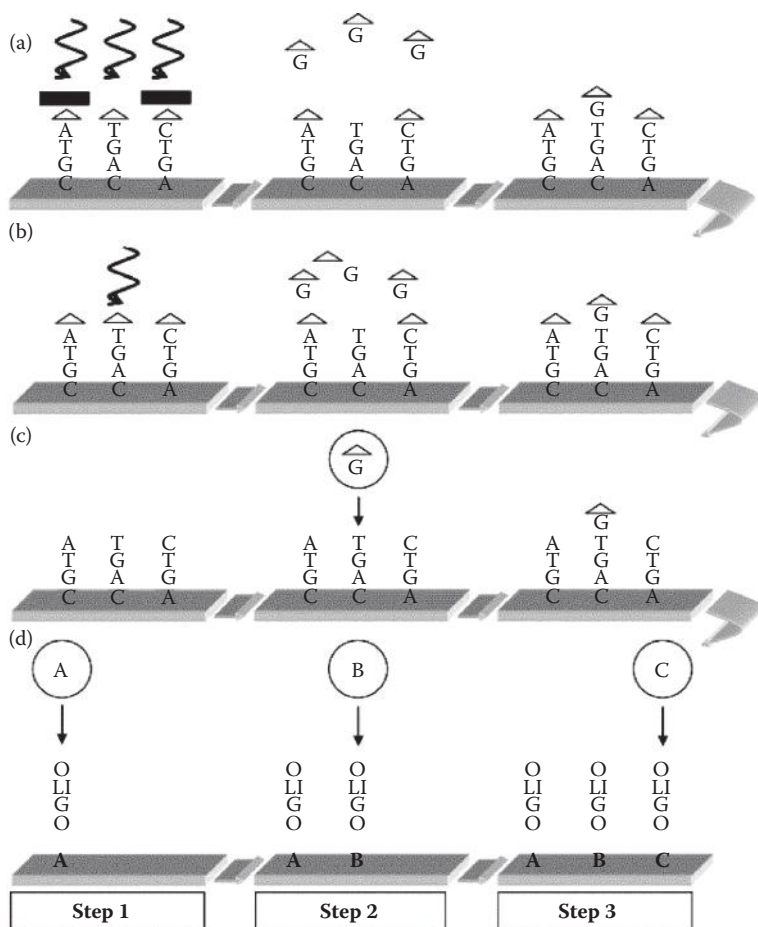
the use of fluorescent markers Cy3 and Cy5. The labeled RNA “targets” are then spotted on the DNA microarray probes to which the RNA binds in a process called “hybridization.” Before measuring fluorescence, the DNA microarray is washed to remove any excess or unhybridized targets. Fluorescence can be quantified by confocal laser scanning microscopy or other fluorimeters. Using internal standards, fluorescence can be used to calculate the number of targets that hybridize to the probes, which can then be used to determine the change in gene expression in a normal state compared to an experimental state. Running a DNA microarray experiment is relatively quick and easy; the biggest factor detracting from more widespread use is the lengthy fabrication process.

### 5.10.2 Array Fabrication

There are three primary methods for DNA microarray fabrication: in situ synthesis, contact printing, and non-contact printing (Dufva et al., 2005). In situ synthesis creates the oligonucleotides directly on the microarray and is capable of producing the highest quality microarrays, although these high-quality arrays require expensive robotics systems and the use of clean room techniques. All methods of in situ synthesis use nucleosides that have already been modified with a protecting group to ensure that the correct sequences of oligonucleotides are present upon completion. In the first method [Figure 5.35(a)], nucleosides are modified with the photolabile reagent 2-nitrophenyl propoxycarbonyl (NPPOC). These groups are removed using light and can then react with available nucleosides. Synthesis is controlled by the use of various masks that are placed on top of the microarray, allowing only specific regions to react. After subjecting a mask to light, the appropriate nucleoside is added, the microarray is washed, and then the process repeats. The downside to this method is that it requires a different mask for each step of the synthesis. Another method [Figure 5.35(b)] involves similar chemistry, yet instead of masks, a series of mirrors is used to direct light to the spots to be deprotected. Yields of 77% for 25 base pair oligonucleotides have been achieved using NPPOC as the protecting group.

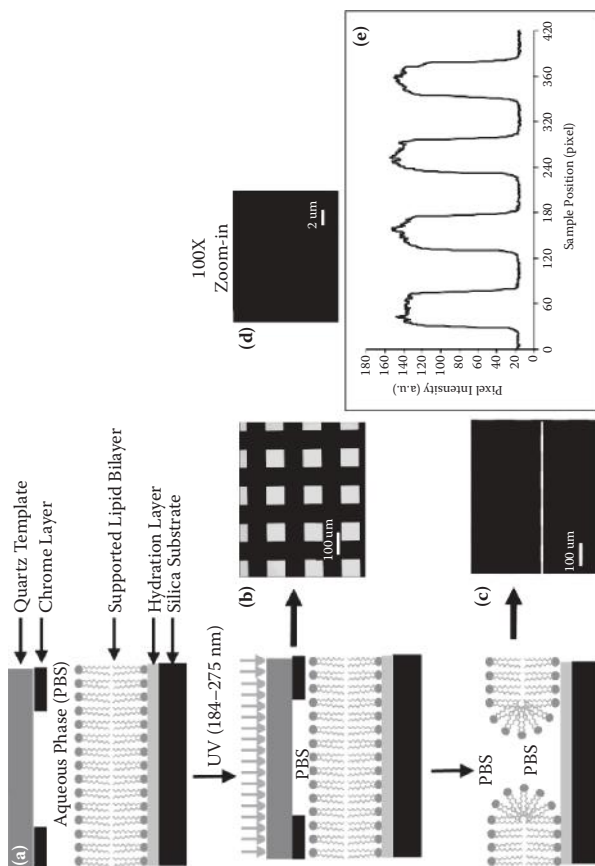
Because they are cheap and still feasible, printing methods are generally the preferred methods of DNA microarray fabrication. Contact printing uses a robotic system of high definition pins to dispense quantities of probe solutions at given coordinates. Non-contact printing differs in that instead of pins, inkjets or similar technologies are used to dispense





**Figure 5.35** Methods of fabrication. (a) In situ with masks. (b) In situ with mirrors. (c) In situ printing. (d) Printing pre-synthesized oligonucleotides. (Reprinted from Dufva, M. Fabrication of High Quality Microarrays. *Biomolecular Engineering* 2005, 22, 173–184. Copyright 2006, with permission from Elsevier.)

picoliter quantities of solution. Non-contact printing is preferred over contact printing as non-contact printing creates smaller spots, and thus more spots can be fit on a microarray. Non-contact printing has been used in one method of in situ synthesis [Figure 5.35(c)]. In this method, all bases are deprotected followed by site-specific dispensing of phosphoramidite-protected nucleosides; cycles of deprotection and the addition of nucleosides are repeated to complete oligonucleotide synthesis. Perhaps the most practical methods for fabrication of DNA microarrays are the printing methods in which all oligonucleotides are synthesized beforehand. Via contact or non-contact



**Figure 5.36** Direct patterning of void arrays within bilayer membranes using UV photolithography. (a) A schematic diagram of the key process steps. (b) Bright-field image, where bright squares are quartz and dark regions reveal the chrome background. (c) Epifluorescence images revealing resultant fluorescence patterns. (d) A high-magnification fluorescence image of the sharp boundary between the UV-exposed and UV-protected regions. (e) Fluorescence intensity profile across an arbitrarily chosen line spanning four alternating UV-exposed and unilluminated bilayer regions. (Reprinted with permission from Yee, C. K., Amweg, M. L., and Parikh, A. N. Direct Photochemical Patterning and Refunctionalization of Supported Phospholipid Bilayers. *Journal of the American Chemical Society* 2004, 126.43:13962–13972. © American Chemical Society.)

printing, known quantities of each individual oligonucleotide are dispensed at their desired position [Figure 5.35(d)].

### 5.10.3 Optimization

A great deal of optimization is required before a DNA microarray can be used effectively. The quality of a DNA microarray is greatly influenced by the method in which it was fabricated. For example, spot density, the number of spots that can be placed in a given area of a microarray, varies greatly among in situ, non-contact printing and contact printing methods. Using in situ methods, spot sizes with diameters less than 10  $\mu\text{m}$  have been created, whereas printing methods can create spot sizes between 20 and 30  $\mu\text{m}$ . Also affecting spot density is the array geometry. It is important for spots to be arranged in a geometry that maximizes spot density while preventing any overlap; additionally, precise array geometry is significant for future data analysis. Currently, our ability to create high spot density arrays has surpassed our ability to accurately interpret fluorescence. Spot sizes of approximately 30  $\mu\text{m}$  are the smallest that result in distinct fluorescence from separate spots. The creation of higher resolution fluorimeters will be required for further miniaturization of DNA microarrays.

Spot morphology is another parameter that must be monitored. A homogeneous spot is preferred for data analysis. Poor spot morphology is less of an issue with light-directed in situ syntheses. Spot morphology of printing methods can be improved by controlling humidity and temperature, as well as by adjusting the spotting buffer.

Perhaps the most important factors to be controlled for optimization are probe density and hybridized density. These two factors are closely related, as probe density is the number of probes in a given spot, and hybridized density is the fraction of these probes that hybridize when saturated with target molecules. In order to get the greatest fluorescent signal, hybridized density must be maximized; however, a number of issues arise. If probe density is too high, probe molecules may interact with other probe molecules on the surface and become unable to hybridize. Additionally, high probe density leads to high steric hindrance, which prevents targets from hybridizing. A number of troubleshooting methods have been developed to optimize these parameters. In many cases, the microarray surface and the probe molecules can be modified to maximize functional probe density.

Another method to optimize hybridized density uses linker molecules to move the probes farther from the surface to

minimize or prevent any unwanted interactions. Innovation of the linker has resulted in a dendrimeric linker, which has been shown to greatly increase probe density and hybridized density.

#### 5.10.4 Applications

DNA microarrays are excellent tools for studying gene expression and have valuable uses in both research and industry. One of the first uses of DNA microarrays was in the study of *E. coli*. Samples of stationary-phase *E. coli* and log-phase *E. coli* were analyzed by DNA microarray. The data found from microarray analysis led to the discovery of multiple growth-regulating genes and an overall improved understanding of this significant microorganism (Ye et al., 2001).

Industrially, DNA microarrays are of great significance in optimizing biocatalysis. Particularly, fermentation processes are often optimized using DNA microarrays. As biological hydrogen fuel cells gain more attention, it is likely that DNA microarrays will be of great importance in optimizing these systems.

DNA microarrays have great potential to impact healthcare. Intelligent drug design has benefited from the use of DNA microarrays as it allows drug developers to anticipate adverse effects or, alternatively, to discover new drug targets. Other kinds of microarrays are also effective in diagnosing diseases. DNA microarrays have been designed to screen for inherited genetic diseases; also, changes in the expression of certain genes can be attributed to cancer and other chronic illnesses, allowing for earlier detection. Perhaps the most exciting possibility with DNA microarrays is the development of personalized medicine. Using a DNA microarray to better understand an individual's genotype will allow doctors to more effectively treat each person, ultimately resulting in a healthier country.

#### 5.10.5 Arrays of Supported Bilayers and Microfluidic Platforms

Spatially addressed microarrays have found use as extremely rapid and powerful means of data collection. Unlike DNA-, protein-, or peptide-based arrays, phospholipid bilayer systems must remain hydrated at all times in order to retain the desired supramolecular structure (Figure 5.36). This requirement creates a significant challenge for creating arrays of supported bilayers.

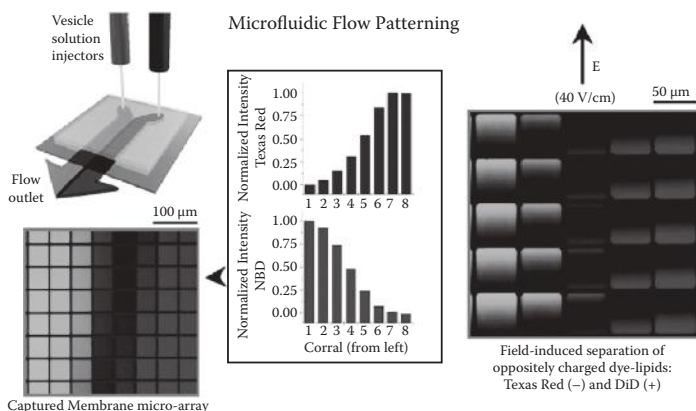
The first method for patterning surfaces with solid supported phospholipid bilayers was created by Groves et al. (1997). A typical formation procedure involved the patterning of photoresist on fused quartz wafers by means of standard

photolithographic techniques. SUVs were then fused onto the substrate between the barriers, creating a lithographically patterned array of identical planar-supported membranes.

Arrays of supported membranes can also be fabricated by selectively destroying regions of a continuous supported bilayer. This is achieved by high-intensity UV illumination through a photomask under aqueous conditions (Yee et al., 2004). The UV radiation generates both ozone and singlet oxygen in highly localized regions. These species react with and degrade the lipids to form water-soluble components. These patterns of holes or of corralled bilayers display long-term stability, retaining their geometric shapes, sizes, and distribution as well as their relative position on the substrate surface.

As shown in Figure 5.37, sharp corners of geometrical features on the mask always result in smoothly curved or rounded edges. This observation is attributed to steric crowding and line-tension effects of the phospholipids. This approach is applicable for producing void patterns of arbitrary shapes, sizes, and densities at predetermined regions within the bilayer.

A neat feature of this patterning method is that the voids can be backfilled by subsequent exposures to the same or a different vesicle solution. This makes it possible to manipulate membrane compositions and dynamically probe lipid-lipid diffusive processes. POPC bilayer samples patterned using the photochemical method shown in Figure 5.37, when exposed to small unilamellar vesicles, causes the nonfluorescent voids to

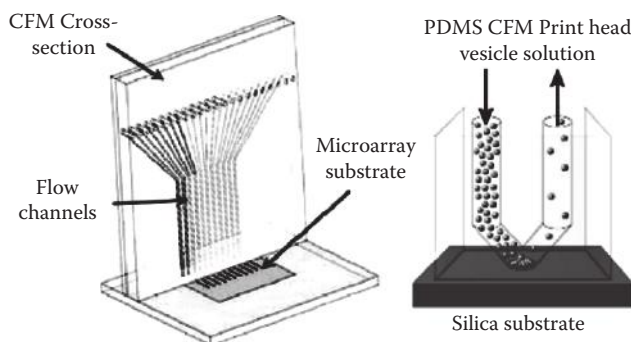


**Figure 5.37** Addressing by laminar flow in a microfluidic channel. Diffusive mixing in a microchannel under laminar flow conditions provides a concentration gradient of different dye-labeled vesicles. (Reprinted with permission from Groves, J. T. and Boxer, S. G. Micropattern Formation in Supported Lipid Membranes. *Acc. Chem. Res.* 2002, 35: 149–157. © American Chemical Society.)

be filled with lipids from the secondary vesicles. Depending on the composition of the secondary vesicles, the pattern can be erased or retained. If the lipids in the secondary vesicles are the same as the initial patterned bilayer, the lipids will diffuse and homogenize quickly, thereby erasing the pattern. If the lipids in the secondary vesicles have significantly different translational mobility, they will retain the backfilled pattern longer. This suggests the possibility of creating a fluid bilayer background with patterned microdomains at specific locations.

The use of laminar flow inside microfluidic channels is also an effective means of producing composition arrays of supported phospholipid bilayers in which two distinct chemical components can be varied simultaneously along a one-dimensional gradient. This allows for the addressing of patterned substrates by the flow of concentration gradients of SUVs formed by the diffusion mixing of two different SUV solutions. Figure 5.38 demonstrates the process of forming a one- or two-component composition array by laminar flow in microfluidic channels. A drawback to this method is the limited number of distinct components that can be simultaneously addressed as well as the lack of control over the positioning of the bilayers.

More recently, Smith et al. (2008) used a 3-D continuous flow microspotter (CFM) system for the preparation of multianalyte lipid bilayer arrays. This method is capable of producing higher density multicomponent arrays compared to traditional 2-D microfluidics. The poly(dimethylsiloxane), or PDMS, microspotter consists of a series of inlet and outlet wells connected by pairs of microfluidic channels embedded within the



**Figure 5.38** Schematic of the CFM apparatus and closeup of the CFM print head in contact with a silica substrate used for bilayer formation. (Reprinted with permission from Smith, K. A., Gale, B. K., and Conboy, J. C. Micropatterned Fluid Lipid Bilayer Arrays Created Using a Continuous Flow Microspotter. *Analytical Chemistry* 2008, 80.21: 7980–7987. © American Chemical Society.)

polymer. When the PDMS print head contacts the substrate, one continuous channel is formed between the inlet and outlet pairs, resulting in the continuous flow of solution over the substrate. Each channel is individually addressable, allowing for the production of 2-D bilayer arrays. A prepatterned substrate is not necessary because the bilayers are effectively corralled into discrete micrometer-sized domains by the residual PDMS deposited on the silica substrate from the PDMS print head, which prevents the lipids from spreading. The packing of the lipids within a  $400 \times 400\text{-}\mu\text{m}^2$  area with spacing of  $400\text{ }\mu\text{m}$  between areas resulted in well-behaving bilayers.

### 5.10.6 Summary

The DNA and lipid microarrays are fairly new technologies whose potential utility has yet to be reached. Widespread use is currently limited by the length of time involved in oligonucleotide synthesis and array fabrication techniques, which make the cost of a microarray experiment rise significantly. As technology advances, microarrays are likely to be an essential tool in any biochemist's or molecular biologist's arsenal.

### CITED REFERENCES

- Baddour, C. E.; Breins, C. Carbon Nanotube Synthesis: A Review. *Int. J. Chem. React. Eng.* 2005, 3, 1–20.
- Blodgett, K. B. Films Built by Depositing Successive Unimolecular Layers on a Solid Surface. *J. Am. Chem. Soc.* 1935, 57, 1007–1022.
- Castellana, E. T.; Cremer, P. S. Solid Supported Lipid Bilayers: From Biophysical Studies to Sensor Design. *Surf. Sci. Rep.* 2006, 61(10), 429–444.
- Crossland, E. J.; Ludwigs, S.; Hillmyer, M. A.; Steiner, U. Freestanding Nanowire Arrays from Soft-Etch Block Copolymer Templates. *Soft Matter*. 2007, 3, 94–98.
- Dai, H. Carbon Nanotubes: Opportunities and Challenges. *Surf. Sci.* 2002, 500, 218–241.
- Drexhage, K. H.; Zwick, M. M.; Kuhn, H. Sensitized Fluorescence after Light-Released Energy Transfer Through Thin Layers. *Ber. Bunsen-Ges. Phys. Chem.* 1963, 67, 6267.
- Dufva, M. Fabrication of High Quality Microarrays. *Biomol. Eng.* 2005, 22, 173–184.
- Gao, X.; Yang, L.; Petros, J. A.; Marshall, F. F.; Simons, J. W.; Nie, S. *In Vivo* Molecular and Cellular Imaging with Quantum Dots. *Curr. Opin. Biotechnol.* 2005, 16, 63–72.

- Groves, J. T.; Boxer, S. G. Micropattern Formation in Supported Lipid Membranes. *Acc. Chem. Res.* 2002, *35*, 149–157.
- Groves, J. T.; Ulman, N.; Boxer, S. G. Micropatterning Fluid Lipid Bilayers on Solid Supports. *Science*. 1997, *275*(5300), 651–653.
- Gu, L. Q., et al. Stochastic Sensing of Organic Analytes by a Pore-Forming Protein Containing a Molecular Adapter. *Nature*. 1999, *398*(6729), 686–690.
- Johal, M. S.; Parikh, A. N.; Lee, Y.; Casson, J. L.; Foster, L.; Swanson, B. I.; McBranch, D. W.; Li, D. Q.; Robinson, J. M. Study of the Conformational Structure and Cluster Formation in a Langmuir-Blodgett Film Using Second Harmonic Microscopy and FTIR Spectroscopy. *Langmuir* 1999, *15*, 1275.
- Keidar, M.; Waas, A. M. On the Conditions of Carbon Nanotube Growth in the Arc Discharge. *Nanotechnology* 2004, *15*, 1571.
- Klitzing, K. V.; Dorda, G.; Pepper, M. New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance. *Phys. Rev. Online Arch.* 1980, *45*, 494–497.
- Knobler, C. M. *Science* 1990, *249*, 870.
- Lei, G.; MacDonald, R. Lipid Bilayer Vesicle Fusion: Intermediates Captured by High-Speed Microfluorescence Spectroscopy. *Biophysical Journal*, 2003, *85*, 1585–1599, Biophysical Society, Elsevier.
- Möbius, D. Z. *Naturforsch.* 1969, *24a*, 251.
- Radeva, T.; Petkanchin, I. Electrical Properties and Conformation of Polyethylenimine at the Hematite-Aqueous Interface. *J. Colloid Interface Sci.* 1997, *196*, 87.
- Richter, R. P.; Berat, R.; Brisson, A. R. Formation of Solid-Supported Lipid Bilayers: An Integrated View. *Langmuir*. 2006, *22*(8), 3497–3505.
- Saito, R.; Dresselhaus, G.; Dresselhaus, M. S. *Physical Properties of Carbon Nanotubes*. Imperial College Press: London, 1998.
- Sinnot, S. B.; Andrews, R.; Qian, D.; Rao, A. M.; Mao, Z.; Dickey, E. C.; Derbyshire, F. Model of Carbon Nanotube Growth through Chemical Vapor Deposition. *Chem. Phys. Lett.* 1999, *315*, 25–30.
- Smith, K. A.; Gale, B. K.; Conboy, J. C. Micropatterned Fluid Lipid Bilayer Arrays Created using a Continuous Flow Microspotter. *Anal. Chem.* 2008, *80*(21), 7980–7987.
- Takayanagi, K. Suspended Gold Nanowires: Ballistic Transport of Electrons. *JSAP Int.* 2001, *3*, 3–8.



- Tamm, L. K.; McConnell, H. M. Supported Phospholipid-Bilayers. *Biophys. J.* 1985, 47(1), 105–113.
- Thostenson, E. T.; Resn, Z.; Chou, T. Advances in the Science and Technology of Carbon Nanotubes and Their Composites: A Review. *Compos. Sci. Technol.* 2001, 61, 1899–1912.
- Wang, et al. AFM Tip Hammering Nanolithography, *Small* 2009, 5(4), 477–483.
- Ye, R.; Wang, T.; Bedzyk, L.; Croker, K. Applications of DNA Microarrays in Microbial Systems. *J. Microbiol. Meth.* 2001, 47, 257–272.
- Yee, C. K.; Amweg, M. L.; Parikh, A. N. Direct Photochemical Patterning and Refunctionalization of Supported Phospholipid Bilayers. *J. Am. Chem. Soc.* 2004, 126(43), 13962–13972.

## REFERENCES AND RECOMMENDED READING

- Hamley, I. W. 2007. *Introduction to Soft Matter*, Revised Edition. John Wiley & Sons, Chichester, West Sussex, UK. Chapter 6 provides an excellent read on lipid membranes, proteins, and other macromolecular assemblies. Since material on liquid crystals is not covered in detail, it is recommended that the student read Chapter 5 in this book. This book also has some excellent problems.
- Kuhn, H., and Försterling, H.-D. 2000. *Principles of Physical Chemistry: Understanding Molecules, Molecular Assemblies, Supramolecular Machines*. John Wiley & Sons, Chichester, West Sussex, UK. Chapters 22 (“Organized Molecules Assemblies”) and 23 (“Supramolecular Machines”) are crucial reads for anyone interested in the physical chemistry of self-assembly and supramolecular processes. These chapters are very well written and accessible to undergraduates.
- Hanson, G. W. 2008. *Fundamentals of Nanoelectronics*. Prentice-Hall, Upper Saddle River, NJ. This book gives a nice quantum mechanical treatment of nanoelectronics. Chapter 9, “Nanowires and Nanotubes” is particularly useful. This book assumes a strong background in physics and is recommended for students interested in free and confined electrons in nanomaterials.
- Lehn, J. M. 1995. *Supramolecular Chemistry*. VCH Weinheim. An essential reference to supramolecular

chemistry written by the Nobel Laureate who coined the term.

- Prasad, P. N. 2003. *Introduction to Biophotonics*. John Wiley & Sons, Hoboken, NJ. This book, while not focusing on nanomaterials, provides some very interesting examples of biomaterials that are used in nanofilms. These include materials for biosensors and microarray technology for genomics and proteomics. There is also an interesting chapter on bionanophotonics.
- Prasad, P. N. 2004. *Nanophotonics*. John Wiley & Sons, Hoboken, NJ. This is a highly recommended read. There is excellent coverage on nanolithography and nanophotonics for biotechnology and nanomedicine.
- Decher, G., and Schlenoff J. B. (Eds.). 2003. *Multilayer Thin Films: Sequential Assembly of Nanocomposite Materials*. John Wiley & Sons, Weinheim, Germany. This is the best source of information on electrostatic self-assembly of polyelectrolytes. This book is a crucial read for those interested in the layer-by-layer assembly of polyelectrolytes. The book provides excellent coverage of both fundamental principles and potential applications of polyelectrolyte multilayers.
- Rao, C. N. R., Müller, A., and Cheetham, A. K. (Eds.). 2005. *The Chemistry of Nanomaterials: Synthesis, Properties, and Applications (Volumes 1 and 2)*. John Wiley & Sons, Weinheim, Germany. This book contains some of the best reviews on quantum dots, nanotubes, and nanowires (synthesis and properties). The book also contains a good chapter on oxide nanoparticles.
- Rao, C. N. R., Müller, A., Cheetham, A. K. (Eds.). 2007. *Nanomaterials Chemistry: Recent Developments and New Directions*. John Wiley & Sons, Weinheim, Germany. This book provides good coverage of mostly inorganic nanomaterials. The book describes the use of nanomaterials for some interesting applications such as supercapacitors, molecular machines, and transistors.
- Gompper, G., and Schick, M. (Eds.). 2006. *Soft Matter*. John Wiley & Sons, Weinheim, Germany. The first volume in this book explores polymer melts and mixtures. The second volume focuses on complex colloidal suspensions. Both volumes are heavy on theoretical studies and are recommended for graduate students seriously interested in computational chemistry.

## END OF CHAPTER QUESTIONS

1. (a) Using antibodies A and D, which bind specifically to an antigen X at different sites, and a UV-vis spectrophotometer, design a series of experiments to quantify the concentration of X in solution. Explain all steps/calculations.  
(b) What kind of repulsive and attractive interactions would you expect between two zwitterionic lipid bilayers? Between a zwitterionic lipid bilayer and solid support? How are they different?  
(c) Consider a situation where quantum dots are injected into a cell for imaging—will the size of quantum dots interfere with cell functions or pose any problems in this experiment?
2. Consider the three molecules shown in Question 11, Chapter 2. Langmuir films that can be constructed from these molecules show hysteresis in their  $\Pi$ -A isotherm. Hysteresis occurs when the compression isotherm differs from the expansion isotherm. Hysteresis in the isotherm for these molecules is largest for the one with the strongest end group dipole moment. Explain this observation.
3. Using the following materials and your knowledge of self-assembled monolayers, design a way to immobilize an antibody (containing primary amines) on a gold surface. The surface is functionalized through a process of incubation in the given solutions, and QCM is then used for the immobilization of the antibody itself. Give details as to why and how each material binds to the underlying layer. Qualitatively predict what the QCM mass versus time profiles will look like for each deposition step. Materials: (a) An aqueous solution of 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (or EDC), (b) mercaptoundecanoic acid in anhydrous ethanol, (c) an aqueous solution of N-hydroxysuccinimide (NHS), (c) a gold-coated substrate.
4. Consider the fabrication of a polyelectrolyte multilayer composed of PEI and PAZO on a glass substrate. After each deposition step, the mass of the film can be measured by QCM and thickness by ellipsometry. Predict how these values change for a 10-bilayer film as the following parameters are varied: (a) Decreasing the pH

of the polyelectrolyte solution from 7 to 4, (b) increasing the salt concentration from 0 to 100 mM, and (c) increasing the deposition temperature from 25°C to 30°C. For (a) you will need to look up various  $\text{pK}_a$  values for groups on each polyelectrolyte.



# Appendix

---

## TOTAL INTERNAL REFLECTION (TIR)

When light is shone upon the boundary between two mediums or substances of different refractive indexes ( $n_1 \neq n_2$ ), the light passes through the boundary into the second medium and is refracted toward the normal if  $n_2 > n_1$  and away from the normal if  $n_1 > n_2$ . This refraction obeys Snell's law:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad (\text{A.1})$$

where  $n_1$  and  $n_2$  are the refractive indexes of the two substances,  $\theta_1$  is the angle of incidence, and  $\theta_2$  is the angle of refraction. In cases where the refractive index of the second medium is less than that of the first (i.e.,  $n_1 > n_2$ ), there exists an angle, called the critical angle  $\theta_{\text{critical}}$ , where the angle of refraction is  $90^\circ$ , or where the light in the second substance is refracted along the interface between the two substances. The critical angle  $\theta_{\text{critical}}$  between any two substances can be calculated by applying Snell's law and performing some basic algebra to get

$$\theta_{\text{critical}} = \arcsin\left(\frac{n_2}{n_1}\right) \quad (\text{A.2})$$

When incident light strikes the interface between the substances at an angle that is greater than the critical angle, the light is reflected from the surface. This phenomenon is called total internal reflection.

## EVANESCENT WAVES

A description of total internal reflection using classical physics says that the incident light is completely reflected from the surface between the two substances. In reality, however, some of the energy from the incident light actually penetrates the second medium to a small extent. This “portion” of the light that enters the other medium is called an evanescent wave. The evanescent wave (or evanescent field) decays rapidly, moving

away from the interface between the two substances according to the following equation:

$$E_x = E_0 e^{-x/d_p} \quad (\text{A.3})$$

where  $E_x$  is the electric field amplitude of the evanescent wave at a distance  $x$  from the interface,  $E_0$  is the electric field at the interface, and  $d_p$  is the penetration depth, defined as the distance at which  $E_0$  is reduced to  $1/e$  of its original value. If the conditions of TIR that generate an evanescent wave are known, then  $d_p$  can be calculated as

$$d_p = \frac{\lambda}{2\pi n_1 \sqrt{\sin^2 \theta_{\text{incidence}} - \left(\frac{n_2}{n_1}\right)^2}} \quad (\text{A.4})$$

where  $\theta_{\text{incidence}}$  is the angle of incidence, and  $n_1$  and  $n_2$  are the refractive indexes of the two mediums. We see that the penetration depth  $d_p$  of the evanescent wave can be calculated for a given wavelength of light. In the visible region, the values of  $d_p$  typically range from 50 to 100 nm. Thus, the evanescent wave can be an excellent probe of the area near the boundary between the two substances (i.e., it is an excellent probe of surface modifications), and as such serves as the basis for a variety of nanomaterial characterization techniques such as SPR, ATR-FTIR, and DPI.

# Glossary

---

**Adsorbate:** A term used to describe an atom, ion, or molecule that is either deposited onto a surface or is the species to be adsorbed.

**Adsorbent:** Surface upon which an adsorbate is deposited.

**Adsorption isotherm:** A plot of how the amount of adsorbate on a given adsorbent changes with pressure (for gas-phase deposition) or concentration (for liquid-phase deposition).

**Aggregation:** Buildup of material at an interface. The term also describes the formation of clusters or an increased local concentration of molecules (e.g., micelle) in bulk phases.

**Aggregation number:** Number of molecules that comprise a single micelle.

**Amphiphile:** Molecule (typically a surfactant) containing both hydrophilic and hydrophobic domains.

**Anisotropic:** Something in which the refractive index depends on the orientation of the propagating wave. The term also refers to an assembly in which all of the molecular components point in the same direction.

**Anti-Stokes shift:** (a) Like the Stokes shift, this is the difference between absorbed and emitted photons. However, in this case the emitted photon must exceed the energy of the exciting photon. This extra energy stems from the dissipation of thermal energy within crystalline structures.

**Anti-Stokes shift:** (b) The frequency shift of Raman scattered light to a higher energy.

**Atomic force microscopy (AFM):** A sensitive imaging technique, which exploits a small piezoelectric tip monitored by a laser to detect minute changes in surface topology.

**Attenuation:** The gradual loss in intensity due to flux through a medium.

**Ballistic transport:** Refers to a nanostructure whose mean-free path is less than its width. Ballistic transporters (typically metal nanowires) have very high conductance values compared to those whose wire widths fail to exceed the mean-free path.



- BET isotherm:** An adsorption isotherm that describes multi-layer film formation. This isotherm assumes that physisorption on the substrate is infinite, that no interlayer interactions exist, and that each layer can be described by the Langmuir model.
- Black lipid membrane:** Lipid membrane assembled on an aperture in a hydrophobic surface. The membrane is referred to as “black” because it is dark when exposed to light as photons reflected off the back half of the bilayer destructively interfere with those bouncing off the front half.
- Bottom-up synthesis:** In this approach, nanostructures are built by continuously extending a thread of bound molecules. Subsequent layers are built atop previous ones exploiting intermolecular forces.
- Bulk phase:** Solvent region above or below an interface. The term is used in context to adsorption. For example, molecules move from the bulk phase and aggregate on the surface.
- Cassie-Baxter wetting:** A wetting state in which water rests upon nano- or micro-sized “pins” with air spaces in between.
- Charge overcompensation:** Refers to the tendency of a charged layer to attract excess counter ions during self-assembly. In this way, layers can be subsequently built atop one another with the overcompensated charge acting as a medium for continued growth of molecules of alternating charge.
- Charge reversal:** Integral to the iterative process of electrostatic self-assembly. A charged molecule can aggregate atop a preexisting surface of opposite charge. The resulting surface is the opposite charge of the original layer, an instance of charge reversal. See also charge overcompensation.
- Chemisorption:** Process of using chemical bonds to tether molecules to a surface.
- Chromophore:** The region of a molecule that absorbs light and undergoes an electronic transition.
- Column chromatography:** Method of purifying and separating one chemical compound from another.
- Conjugation:** Alternating single and double bonds in linear carbon chains. All carbon atoms are  $sp^2$  hybridized and the electrons in the  $\pi$ -bonds are actually delocalized along the entire carbon chain. See also pi-bonding.

**Constructive interference:** A phenomenon in which light waves combine to produce a more intense wave, or brighter light area, than the individual waves. This type of interference occurs when the waves are in-phase.

**Contact angle:** Angle formed between a drop of liquid and a solid interface.

**Contact angle hysteresis:** The difference between a droplet's receding and advancing contact angle.

**Converse piezoelectric effect:** The ability of a crystal to oscillate due to applied alternating current.

**Cooperative adsorption:** An adsorption processes whereby the presence of one adsorbate enhances the adsorption of another. Also refers to the property set of positively and negatively charged molecules, which do not deposit on a surface individually, to stabilize electrostatic interactions and allow for the deposition of both molecules on a surface simultaneously.

**Critical micelle concentration (CMC):** Concentration of surfactant molecules in solution above, in which micelles begin to spontaneously form.

**Debye length:** The characteristic decay length of the electrostatic potential for the Debye-Hückel model. Also used as a rough approximation for the size of the electric double layer.

**Desorption:** Process of a molecule bound to a surface becoming detached from that surface and returning to the bulk phase.

**Destructive interference:** A phenomenon in which light waves combine to produce a less intense wave, or dimmer light area, than the individual waves. This type of interference occurs when the waves are completely out-of-phase.

**Direct piezoelectric effect:** The ability of a crystal to produce an electric field in response to mechanical stress.

**Dispersion forces:** Force between two instantaneously induced dipoles.

**Dissipation:** Value for relating the energy lost to the energy stored over the course of an event.

**DNA microarray:** DNA microarrays are surfaces that have been coated with specific oligonucleotide sequences and have been particularly useful in gene expression studies. DNA microarrays vary in the number of oligonucleotides as is determined by the nature of the study; diagnostic DNA microarrays generally use tens

of oligonucleotides, whereas those for research and screening can have hundreds of thousands of oligonucleotides on a single microarray.

**Electrical double layer:** The diffuse layer of counter ions in a solution, which are associated with a charged surface.

**Evanescent field:** Standing waves formed at the boundary between two media with different wave motion properties. They decay exponentially with distance from the boundary in which the waves are formed.

**Exciton:** A quasiparticle created by irradiation of photoactive materials and consisting of an unpaired electron and an imaginary “electron hole.” As incident light is exposed to a dye interface, an in-phase oscillation among all the layer’s oscillators occurs and an excited domain (the exciton) forms, having absorbed the energy from the incident photons.

**Extinction coefficient:** Also called molar absorptivity, it is a parameter that helps define how strongly a substance absorbs light. It is often given at a given wavelength per mass unit or per molar concentration.

**Fluorescence:** A relaxation process in which the molecule relaxes by reemitting light, generally of lower energy than the light it absorbed. More specifically, it refers to the light emitted when an electronic transition occurs between electronic states of the same spin multiplicity (e.g., singlet-singlet transitions).

**Fluorescence interference contrast microscopy (FLIC):** Monitors the calculable modulation of fluorescence intensity due to the interaction between a reflecting surface and fluorescent objects in order to attain nanometer-accurate height measurements. In this method, a film on a reflective surface is capped with a fluorescently tagged entity and the assembly’s size is quantified based on the specific interaction between the surface and fluorophore.

**Fluorescence resonance energy transfer (FRET):** Mechanism of energy transfer between chromophores whereby an excited donor transfers an electron over a short (<10nm) distance to an acceptor molecule through dipole–dipole coupling.

**Fluorophore:** Fluorescent unit of a molecule, which absorbs and emits energy at a specific wavelength.

**Fractal:** Geometric shape that can be split into sections that are at least partially identical to a smaller replica of the entire structure.

**Free electron model:** In this model, the electron does not exist as a discrete particle moving along the line. Rather, it resembles a standing wave whose exact form depends on the value of  $n$ .

**Gravimetry:** The measure of the strength of gravitational fields.

**H-type aggregate:** One-dimensional molecular assembly in which the dipole moments are aligned parallel to each other but perpendicular to the line joining their centers. This is sometimes referred to as the “face-to-face arrangement.”

**Hard sphere model:** A way to determine atomic radius by assuming atoms in a solid are hard spheres and pack closely together.

**HOMO:** Highest occupied molecular orbital.

**Hydrodynamic radius:** The radius of a sphere that diffuses at the same rate as the molecule. Because most molecules are not spherical, this radius is often smaller than the effective rotational radius.

**Hydrogen bonding:** Attractive interaction between a hydrogen atom and another electronegative atom.

**Hydrophobic effect:** Tendency of non-polar molecules to aggregate in polar solvents to reach a thermodynamically favored energy state.

**Induced dipole interactions:** Force existing between a permanent dipole and a neighboring induced dipole.

**Interface:** The two-dimensional region of space at which two different phases contact each other.

**Interferometry:** The study of the ways in which light waves interact or interfere with each other.

**Intermolecular force (F):** Forces acting between sets of molecules, such as hydrogen bonding or dipole–dipole interactions.

**Interpenetration:** In polyelectrolyte multilayer films, this is the tendency of polycation and polyanion to commingle to form highly homogeneous assemblies rather than distinct, stratified layers. This process shields the excess charge within the distinct layers, allowing for tighter packing.

**Ion-ion forces:** Attractive or repulsive interaction between ionic species.

**J-type aggregate:** One-dimensional molecular assembly in which the dipole moments of the individual monomers are aligned parallel to the line joining their centers. This is sometimes referred to as the “end-to-end arrangement.”

**Laminar flow:** The flow of fluid in parallel layers.

**Langmuir adsorption isotherm:** Equation relating the concentration or pressure of an adsorbate to the degree of deposition on an adsorbent.

**Langmuir-Blodgett deposition:** The transfer of a monolayer to a substrate using a trough, which compresses molecules (often surfactants) on the surface of a liquid buffer. The hydrophobic tails of the molecules will assemble on the surface of the substrate as it is passed through the compressed layer of surfactant. Multilayer films can be created by switching direction of the substrate entering the compressed layer.

**Layer-by-layer deposition:** Process of building nanoscale assemblies by exploiting intermolecular or chemical attraction forces. A common type of layer-by-layer assembly (LbL) is electrostatic self-assembly, where, for example, a positively charged polyelectrolyte assembles spontaneously on an oxidized silica or silanol surface.

**Lennard-Jones potential energy curve:** Mathematical approximation describing the interaction between two neutral molecules. The curve describes both attractive and repulsive interactions and the distance at which these interactions cancel each other out.

**Lipopolymer:** Consists of a soft hydrophilic polymer layer with lipidlike molecules at their surface. These structures can insert into a phospholipid membrane and tether to the polymer spacing.

**Lithography:** *See* nanolithography.

**LUMO:** Lowest unoccupied molecular orbital.

**Mean-free path:** Relates to the distance an electron travels between subsequent collisions with other moving particles.

**Micelle:** Spherical structures formed from the aggregation of surfactant molecules above the critical micelle concentration (CMC).

**Microemulsion:** A clear solution of liquid water, oil, and surfactant without phase separation of hydrophobic and hydrophilic entities.

**Molecular Orbital (MO):** Mathematical function that describes the wavelike behavior of electrons in a molecule.

**Molecular polarizability:** The degree of electron density distortion due to an electric field.

**Nanowire:** A nanostructure that has a diameter on the scale of a nanometer ( $10^{-9}$  m) and an unrestricted length.

**Nanolithography:** Refers to a wide variety of nanoscale surface manufacturing techniques for creating patterned surfaces, such as those found in semiconductor circuits.

**Nonlinear optical:** A term that indicates the dielectric polarization of the media responds in a nonlinear manner to the electric field of light.

**Oligonucleotide:** A short nucleic acid polymer comprised of fewer than 20 bases.

**Overlap repulsion:** The interaction that results from the fact that two atoms or molecules cannot occupy the same point in space.

**P-polarized light:** Electric field vector polarized along the plane perpendicular to the plane of incidence.

**Path length:** The distance in which light moves through a material.

**Pauli Exclusion Principle:** States that no two electrons can have the same four quantum numbers.

**Phonon:** A quantum mechanical description of a special type of vibration, known as the normal modes in classical physics. Phonons are the discrete amounts of energy crystalline structures that these modes of vibration can absorb. In this aspect of quantized energy, the phonon is analogous to the photon.

**Phosphorescence:** A relaxation process when an excited electron undergoes nonradiative intersystem-crossing to a slightly more stable excited state, called a triplet state, before relaxing and releasing a photon of light.

**Photobleaching (FRAP):** Photobleaching is the process of destroying a fluorophore with intense light. FRAP, or Fluorescence Recovery After Photobleaching, is the method quantifying lateral diffusion rates by monitoring fluorescence under a microscope and measuring the time it takes for a photobleached segment of a fluorophore film to reorganize. Fluorophores reorganize, intact ones replacing the inactivated ones. The hole in fluorescence created by photobleaching disappears at a rate determined by the lateral diffusion.

**Photorelaxation:** Process of an excited state chromophore dissipating energy through release of a photon.

**Physisorption:** Process of exploiting weaker molecular forces to deposit molecules on the surface. Generally weaker and less specific than chemisorption.

**Pi bonding (or  $\pi$ -bonding):** Bond resulting from the overlap of atomic orbitals that are in contact through two areas of overlap above and below the internuclear axis. For

example, in  $sp^2$  hybridized carbon atoms, the unhybridized 2p orbitals on neighboring carbons overlap to form the  $\pi$ -bond.

**Pi-pi stacking:** Packing of large, often aromatic molecules whereby favorable overlap across pi systems of different molecules occurs.

**Plasma oscillations:** Rapid oscillations of the electron density in a conducting medium.

**Plasmon:** The smallest unit of plasma oscillations; used for quantization.

**Poisson-Boltzmann equation:** Calculates the actual distribution of the counter ions at equilibrium.

**Polarizability:** Defined as the susceptibility of an atom or molecule's electron cloud to distortion by an external electric field.

**Polyelectrolyte:** Polyelectrolytes are polymers, or chains of molecules, which contain free ions that make them electrically conductive. Soluble in water, polyelectrolytes become charged when in solution, and are often countered by a salt ion of opposite charge. The amount of charge on a polyelectrolyte determines whether it is classified as strong or weak. Strong polyelectrolytes are fully soluble, whereas weak polyelectrolytes, with fractional charge, are only partially soluble.

**Polymer cushioned phospholipid bilayers:** Allow bilayers to effectively be decoupled from the surface by minimizing interactions with the underlying substrate.

**Potential energy (V):** Energy stored within a system due to its position in a force field.

**Quantum dots:** Quantum dots are inorganic semiconductor nanoparticles, typically 2–10 nm in size. The excitons (electron-hole pairs) of quantum dots are confined to three dimensions. Quantum dots typically consist of a core, shell, and final coat, as shown in Figure 5.10. They are characterized by the nature of each of these layers, their size and aspect ratio, their quantum efficiency in optical materials, and their coercivity in magnetic materials. Properties of an ensemble of quantum dots are additionally determined by particle size distribution and differences in morphology within the ensemble.

**Raman scattering:** The scattering of light in which the light increases or decreases in energy, often due to a change in vibrational energy of the scattering molecule.

**Rayleigh scattering:** The scattering of light without a change in its energy. The degree of scattering is a relationship

between the wavelength of light being scattered and the electric polarizability of the particles doing the scattering.

**Refractive index:** Ratio of the velocity of light in a vacuum to the velocity of light in a medium.

**Resonant frequency:** The frequency at which the local amplitude of oscillation is a maximum.

**S-polarized light:** Electric field vector polarized along the plane of incidence.

**Scanning tunneling microscopy (STM):** A sensitive imaging technique that exploits a small tunneling current between a surface and a tip to detect minute changes in surface topology.

**Second-harmonic generation:** A nonlinear optical effect in which the frequency of light is doubled when passed through a material.

**Self-assembled monolayer (SAM):** A surface coverage that forms when molecules are spontaneously attracted to a functional surface, often by electrostatic or van der Waals forces. SAMs can also result through chemisorption.

**Self-assembly:** The process whereby molecules in the bulk phase spontaneously organize either in the bulk phase or on a surface.

**Sigma bonding:** A strong covalent bond characterized by symmetry with respect to rotation about the bond axis.

**Small unilamellar vesicles (SUVs):** Tiny (50–200 nm), single-walled vesicles of uniform diameter.

**Soft sphere model:** More realistic way of modeling an atom, this model assumes atoms are compressible to some degree and do not have completely rigid boundaries.

**Spectroscopy:** The study of the absorption or transmission of light through a medium as a function of either wavelength or frequency.

**Spot density:** Number of spots that can be placed in a given area of a DNA microarray, varies greatly among in situ, non-contact, and contact printing methods.

**Steric hindrance:** Stress in a molecular structure or assembly that arises when oversized functional groups are forced into too small a space. Such stress can result in torsional strain.

**Stokes shift:** (a) The difference between the excitation photon wavelength necessary to induce an electronically excited state and the wavelength of the subsequently emitted photon. (b) The frequency shift of Raman scattered light to a lower energy.



- Sum-frequency generation:** A nonlinear optical effect when two photons of different frequencies combine to produce a photon of the sum frequency when passed through a material.
- Superhydrophobic surface:** Highly hydrophobic surface with water contact angles exceeding 150°.
- Supramolecular chemistry:** The study of chemical compounds comprised of a distinct number of cooperating molecules. Supramolecular machines are comprised of a “cooperating set” of molecules and exploit intra- and intermolecular forces to complete a specialized task.
- Surface functionalization:** Deliberate attachment of specific molecules to a surface in order to allow for the specific binding of a subsequent adsorbate.
- Surface tension:** The work done in increasing the area of a surface by transporting a molecule from the bulk phase to that surface. A measurement of the cohesion of like molecules at an interface. The tendency of a surface to contract due to energetic factors.
- Surfactant:** Compounds that lower the surface tension of a liquid and form micelles above the critical micelle concentration (CMC).
- Thio-:** A prefix that indicates an oxygen atom in the common compound is instead sulfur. For example, ether has a general structure ROR' and a thioether would have a general structure RSR'.
- Thiol:** Thiols are compounds having the structure RSH (where  $R \neq H$ ).
- Top-down synthesis:** In this approach, structures are synthesized starting from a bulk substrate and chiseled down until only a field of desired assemblies remains.
- Total intermolecular potential energy:** The sum of all the different interactions between two molecules.
- Transducer:** A device that converts one type of energy to another.
- Transmittance:** The ratio of the intensity of a light beam after passing through a sample to the original intensity of the light beam.
- Van der Waals interaction:** Attraction or repulsion between molecules due to interaction with covalent bonds or electrostatic forces.
- Viscoelasticity:** The property of materials to both resist flow when a stress is applied and rapidly return to their original state once the stress is removed.

**Waveguide:** An object that confines and directs waves by means of total reflection of the wave within itself.

**Wenzel wetting:** A wetting state in which water rests upon a surface whose morphology has been altered so that in a given area, water is in contact with more surface than if the surface were completely flat.



**"The writing ... is very fluid. The problems and figures are good. Overall, I learned a great deal...."**

—*Professor Lisa Klein, Rutgers University*

**"I believe the textbook will serve students well in their goal to gain a greater understanding of why nanoscaled systems are of great interest, how they are fabricated, and how they are characterized using a wide variety of analytical instrumentation very commonly found in university and industrial settings."**

—*Professor Marcus D. Lay, University of Georgia*

---

With a selective presentation of topics that makes it accessible for students who have taken introductory university science courses, **Understanding Nanomaterials** is a training tool for the future workforce in nanotech development. This introductory textbook offers insights into the fundamental principles that govern the fabrication, characterization, and application of nanomaterials.

### ***Provides the Background for Fundamental Understanding***

Assuming only a basic level of competency in physics, chemistry, and biology, the author focuses on the needs of the undergraduate curriculum, discussing important processes such as self-assembly, patterning, and nanolithography. His approach limits mathematical rigor in the presentation of key results and proofs, leaving it to the instructor's discretion to add more advanced details or emphasize particular areas of interest.

With its combination of discussion-based instruction and explanation of problem-solving skills, this textbook highlights interdisciplinary theory and enabling tools derived from chemistry, biology, physics, medicine, and engineering. It also includes real-world examples related to energy, the environment, and medicine.

---

Author **Malkiat S. Johal** earned his Ph.D. from the University of Cambridge in England. He later served as a post-doctoral research associate at Los Alamos National Laboratory, New Mexico, where he worked on the nonlinear optical properties of nanoassemblies. Dr. Johal is currently a professor and researcher at Pomona College in Claremont, California. His work focuses on the use of self-assembly and ionic adsorption processes to fabricate nanomaterials for optical and biochemical applications.

 **CRC Press**  
Taylor & Francis Group  
an informa business  
[www.crcpress.com](http://www.crcpress.com)

6000 Broken Sound Parkway, NW  
Suite 300, Boca Raton, FL 33487  
711 Third Avenue  
New York, NY 10017  
2 Park Square, Milton Park  
Abingdon, Oxon OX14 4RN, UK

73109  
ISBN: 978-1-4200-7310-2  
90000  
  
9 781420 073102  
[www.crcpress.com](http://www.crcpress.com)